```
CC       is expressed in all synaptic boutons types, including I, II and
CC       III boutons.
CC   -!- DOMAIN: The Asn-Pro-Phe (NPF) motifs, which are found in proteins
CC       involved in the endocytic pathway, are known to interact with the
CC       EH domain (By similarity).
CC   -!- RNA EDITING: Modified_positions=1186; Note=Partially edited.
CC   -!- MISCELLANEOUS: StnA, which is involved in the same pathway, is
CC       derived from the same dicistronic transcript that encodes these
CC       two different proteins.
CC   -!- SIMILARITY: Belongs to the Stoned B family.
CC   -!- SIMILARITY: Contains 1 SHD (Stonin homology) domain.
CC   -!- SIMILARITY: Contains 1 MHD (Mu homology) domain.
CC   -!- CAUTION: Ref.2 sequence differs from that shown due to erroneous
CC       gene model prediction.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U54982; AAC16666.1; -.
DR   EMBL; AE003200; AAF45320.1; ALT_SEQ.
DR   PIR; T13353; T13353.
DR   FlyBase; FBgn0016975; stnB.
DR   GO; GO:0030139; C:endocytic vesicle; IDA.
DR   GO; GO:0005886; C:plasma membrane; IDA.
DR   GO; GO:0008021; C:synaptic vesicle; IGI.
DR   GO; GO:0005515; F:protein binding; IPI.
DR   GO; GO:0008099; P:synaptic vesicle endocytosis; IMP.
DR   InterPro; IPR001392; Clathrn_med.
DR   Pfam; PF00928; Adap_comp_sub; 1.
KW   Endocytosis; Synapse; Repeat; RNA editing.
FT   DOMAIN      729    903       SHD.
FT   DOMAIN      847   1108       INTERACTION WITH SYT.
FT   DOMAIN      904   1218       MHD.
FT   DOMAIN      211    295       PRO-RICH.
FT   SITE          3      5       NPF 1.
FT   SITE         19     21       NPF 2.
FT   SITE         33     35       NPF 3.
FT   SITE         43     45       NPF 4.
FT   SITE        210    212       NPF 5.
FT   SITE        493    495       NPF 6.
FT   SITE        673    675       NPF 7.
FT   VARIANT    1186   1186       T -> A (in RNA edited version).
FT   CONFLICT    117    117       A -> P (IN REF. 1).
FT   CONFLICT   1012   1012       L -> V (IN REF. 1).
SQ   SEQUENCE   1262 AA;  137768 MW;  2CE67046F8214C81 CRC64;

  Query Match             52.3%;  Score 45;  DB 1;  Length 1262;
  Best Local Similarity   87.5%;  Pred. No. 1e+02;
  Matches    7;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          7 PRPTPPRP 14
              ||| ||||
Db        241 PRPAPPRP 248


RESULT 14
RS7_SYNY3
ID   RS7_SYNY3      STANDARD;      PRT;   156 AA.
AC   P74229;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   30S ribosomal protein S7.
GN   RPSG OR RPS7 OR SLL1097.
OS   Synechocystis sp. (strain PCC 6803).
OC   Bacteria; Cyanobacteria; Chroococcales; Synechocystis.
OX   NCBI_TaxID=1148;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=97061201; PubMed=8905231;
RA   Kaneko T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,
RA   Miyajima N., Hirosawa M., Sugiura M., Sasamoto S., Kimura T.,
RA   Hosouchi T., Matsuno A., Muraki A., Nakazaki N., Naruo K.,
RA   Okumura S., Shimpo S., Takeuchi C., Wada T., Watanabe A.,
RA   Yamada M., Yasuda M., Tabata S.;
RT   "Sequence analysis of the genome of the unicellular cyanobacterium
RT   Synechocystis sp. strain PCC6803. II. Sequence determination of the
RT   entire genome and assignment of potential protein-coding regions.";
RL   DNA Res. 3:109-136(1996).
CC   -!- FUNCTION: One of the primary rRNA binding proteins, it binds
CC       directly to 16S rRNA where it nucleates assembly of the head
CC       domain of the 30S subunit. Is located at the subunit interface
CC       close to the decoding center, probably blocks exit of the E-site
CC       tRNA (By similarity).
CC   -!- SUBUNIT: Part of the 30S ribosomal subunit. Contacts proteins S9
CC       and S11 (By similarity).
CC   -!- SIMILARITY: Belongs to the S7P family of ribosomal proteins.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; D90913; BAA18323.1; -.
DR   PIR; S75864; S75864.
DR   HSSP; P22744; 1HUS.
DR   HAMAP; MF_00480; -; 1.
DR   InterPro; IPR000235; Ribosomal_S7.
DR   InterPro; IPR005717; Ribosomal_S7_b/o.
DR   Pfam; PF00177; Ribosomal_S7; 1.
DR   ProDom; PD000817; Ribosomal_S7; 1.
DR   TIGRFAMs; TIGR01029; rpsG_bact; 1.
DR   PROSITE; PS00052; RIBOSOMAL_S7; 1.
KW   Ribosomal protein; RNA-binding; rRNA-binding; tRNA-binding;
KW   Complete proteome.
SQ   SEQUENCE   156 AA;  17384 MW;  9990887678DDDC6E CRC64;

  Query Match             51.2%;  Score 44;  DB 1;  Length 156;
  Best Local Similarity   46.7%;  Pred. No. 17;
  Matches    7;  Conservative    2;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          2 KGXXLPRPTPPRPIY 16
              :|    || || |:|
Db          4 RGNVKKRPVPPDPVY 18


RESULT 15
ACRL_HUMAN
ID   ACRL_HUMAN     STANDARD;      PRT;   232 AA.
AC   P58840; Q9NU35;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Hypothetical acrosin-like protease (EC 3.4.21.-) (Fragment).
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Blakey S.;
RL   Submitted (JAN-2000) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: Belongs to peptidase family S1.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
```

```
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; AL078621; CAB81647.1; -.
DR   InterPro; IPR009003; Cys_Ser_trypsin.
DR   InterPro; IPR001254; Peptidase_S1.
DR   Pfam; PF00089; trypsin; 1.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; PARTIAL.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
KW   Hypothetical protein; Hydrolase; Serine protease.
FT   NON_TER       1      1
FT   DOMAIN       <1    101       SERINE PROTEASE.
FT   DOMAIN      113    116       POLY-PRO.
FT   DOMAIN      155    181       POLY-PRO.
FT   ACT_SITE     51     51       CHARGE RELAY SYSTEM (BY SIMILARITY).
FT   DISULFID     20     36       BY SIMILARITY.
FT   DISULFID     47     77       BY SIMILARITY.
FT   CARBOHYD     21     21       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   232 AA;  25142 MW;  CF987EB42EBACA7A CRC64;

  Query Match             51.2%;  Score 44;  DB 1;  Length 232;
  Best Local Similarity   87.5%;  Pred. No. 26;
  Matches    7;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          7 PRPTPPRP 14
              ||| ||||
Db        143 PRPLPPRP 150


Search completed: March 11, 2004, 16:57:19
Job time : 26 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: March 11, 2004, 16:53:29 ; Search time 39 Seconds
(without alignments)
145.624 Million cell updates/sec

Title: US-09-980-804-1
Perfect score: 86
Sequence: 1 DKGXXLPRPTPPRPIYXX 18

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_25:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	57	66.3	440	5	Q8IRB3	Q8irb3 drosophila
2	52	60.5	782	11	Q8CGW4	Q8cgw4 mus musculu
3	50	58.1	305	16	Q8YTH2	Q8yth2 anabaena sp
4	50	58.1	671	3	O94113	O94113 pneumocysti
5	49	57.0	377	13	Q7ZX30	Q7zx30 xenopus lae
6	49	57.0	532	10	Q8L462	Q8l462 oryza sativ
7	48	55.8	539	11	Q9DCD5	Q9dcd5 mus musculu
8	48	55.8	539	11	Q8CFL7	Q8cfl7 mus musculu
9	48	55.8	957	16	Q8DK04	Q8dk04 synechococc
10	48	55.8	1013	4	Q9NT81	Q9nt81 homo sapien
11	48	55.8	1015	4	Q8N3X1	Q8n3x1 homo sapien
12	48	55.8	1050	4	Q9Y2L7	Q9y2l7 homo sapien
13	48	55.8	1307	10	Q9LVN1	Q9lvn1 arabidopsis
14	47	54.7	165	16	Q9X9V2	Q9x9v2 streptomyce
15	47	54.7	239	5	Q9VRU9	Q9vru9 drosophila
16	47	54.7	298	4	Q96CP3	Q96cp3 homo sapien
17	47	54.7	392	16	Q82GV6	Q82gv6 streptomyce
18	47	54.7	847	10	Q9XIB6	Q9xib6 arabidopsis
19	47	54.7	862	4	Q9NT23	Q9nt23 homo sapien
20	47	54.7	903	4	O14560	O14560 homo sapien
21	47	54.7	940	10	Q7XTN6	Q7xtn6 oryza sativ
22	47	54.7	949	10	Q9SMA4	Q9sma4 oryza sativ
23	46	53.5	166	6	Q95JQ4	Q95jq4 macaca fasc
24	46	53.5	311	10	Q8H4Z9	Q8h4z9 oryza sativ
25	46	53.5	319	16	Q89P70	Q89p70 bradyrhizob
26	46	53.5	464	12	Q9QEJ3	Q9qej3 cercopithec
27	46	53.5	490	12	Q69023	Q69023 human herpe
28	46	53.5	845	4	Q96H68	Q96h68 homo sapien
29	46	53.5	865	11	Q8VIP2	Q8vip2 rattus norv
30	46	53.5	901	11	Q9EP91	Q9ep91 mus musculu
31	46	53.5	901	11	Q8K120	Q8k120 mus musculu
32	46	53.5	902	4	Q7Z598	Q7z598 homo sapien
33	45	52.3	61	10	Q8LIR4	Q8lir4 oryza sativ
34	45	52.3	96	12	Q9QTW0	Q9qtw0 tt virus. o
35	45	52.3	101	11	Q8BR19	Q8br19 mus musculu
36	45	52.3	185	10	Q9AY89	Q9ay89 oryza sativ
37	45	52.3	203	16	Q9RCX9	Q9rcx9 streptomyce
38	45	52.3	233	12	Q8JN32	Q8jn32 arabis mosa
39	45	52.3	233	12	Q8JN22	Q8jn22 arabis mosa
40	45	52.3	233	12	Q8JN20	Q8jn20 arabis mosa
41	45	52.3	242	10	Q43687	Q43687 vigna ungui
42	45	52.3	256	5	Q9N582	Q9n582 caenorhabdi
43	45	52.3	271	11	Q8CDY2	Q8cdy2 mus musculu
44	45	52.3	296	5	Q9VTQ0	Q9vtq0 drosophila
45	45	52.3	330	11	Q8CIA4	Q8cia4 mus musculu

ALIGNMENTS

RESULT 1
Q8IRB3
ID Q8IRB3 PRELIMINARY; PRT; 440 AA.
AC Q8IRB3;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE CG32241-PA.
GN CG32241.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.H., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabor G.L.,
RA Abril J.F., Agbayani A., An H.J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,

```
RA   Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA   Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA   Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA   Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA   Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA   Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,
RA   Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA   Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA   Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA   Wang Z.Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA   Williams S.M., WoodageT, Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,
RA   Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA   Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA   Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT   "The genome sequence of Drosophila melanogaster.";
RL   Science 287:2185-2195(2000).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Celniker S.E., Adams M.D., Kronmiller B., Wan K.H., Holt R.A.,
RA   Evans C.A., Gocayne J.D., Amanatides P.G., Brandon R.C., Rogers Y.,
RA   Banzon J., An H., Baldwin D., Banzon J., Beeson K.Y., Busam D.A.,
RA   Carlson J.W., Center A., Champe M., Davenport L.B., Dietz S.M.,
RA   Dodson K., Dorsett V., Doup L.E., Doyle C., Dresnek D., Farfan D.,
RA   Ferriera S., Frise E., Galle R.F., Garg N.S., George R.A.,
RA   Gonzalez M., Houck J., Hoskins R.A., Hostin D., Howland T.J.,
RA   Ibegwam C., Jalali M., Kruse D., Li P., Mattei B., Moshrefi A.,
RA   McIntosh T.C., Moy M., Murphy B., Nelson C., Nelson K.A., Nunoo J.,
RA   Pacleb J., Paragas V., Park S., Patel S., Pfeiffer B.,
RA   Phouanenavong S., Pittman G.S., Puri V., Richards S., Scheeler F.,
RA   Stapleton M., Strong R., Svirskas R., Tector C., Tyler D.,
RA   Williams S.M., Zaveri J.S., Smith H.O., Venter J.C., Rubin G.M.;
RT   "Sequencing of Drosophila melanogaster genome.";
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RA   Misra S., Crosby M.A., Matthews B.B., Bayraktaroglu L., Campbell K.,
RA   Hradecky P., Huang Y., Kaminker J.S., Prochnik S.E., Smith C.D.,
RA   Tupy J.L., Bergman C., Berman B., Carlson J.W., Celniker S.E.,
RA   Clamp M., Drysdale R., Emmert D., Frise E., de Grey A., Harris N.,
RA   Kronmiller B., Marshall B., Millburn G., Richter J., Russo S.,
RA   Searle S.M.J., Smith E., Shu S., Smutniak F., Whitfield E.,
RA   Ashburner M., Gelbart W.M., Rubin G.M., Mungall C.J., Lewis S.E.;
RT   "Annotation of Drosophila melanogaster genome.";
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   SEQUENCE FROM N.A.
RA   Adams M.D., Celniker S.E., Gibbs R.A., Rubin G.M., Venter C.J.;
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
RN   [5]
RP   SEQUENCE FROM N.A.
RA   FlyBase;
RL   Submitted (SEP-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AE003481; AAN11603.1; -.
DR   FlyBase; FBgn0052241; CG32241.
DR   InterPro; IPR004019; YLP_motif.
DR   Pfam; PF02757; YLP; 4.
SQ   SEQUENCE   440 AA;  47977 MW;  F0B5B0DDAD834D2E CRC64;

  Query Match             66.3%;  Score 57;  DB 5;  Length 440;
  Best Local Similarity   80.0%;  Pred. No. 1.1;
  Matches    8;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          7 PRPTPPRPIY 16
              |:||||||:|
Db         90 PKPTPPRPVY 99


RESULT 2
Q8CGW4
ID   Q8CGW4      PRELIMINARY;      PRT;   782 AA.
AC   Q8CGW4;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Sox-30.
GN   SOX30.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Tissot C., Bardos J., Freemont P.;
RT   "Characterization of Sox-30 as an rfp interacting protein.";
RL   Submitted (JUL-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY005801; AAF99391.1; -.
DR   MGD; MGI:1341157; Sox30.
DR   GO; GO:0003677; F:DNA binding; IEA.
DR   GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR   InterPro; IPR000910; HMG_12_box.
DR   Pfam; PF00505; HMG_box; 1.
DR   SMART; SM00398; HMG; 1.
DR   PROSITE; PS50118; HMG_BOX_2; 1.
SQ   SEQUENCE   782 AA;  83937 MW;  0D1EEBB17BCB4F41 CRC64;

  Query Match             60.5%;  Score 52;  DB 11;  Length 782;
  Best Local Similarity   88.9%;  Pred. No. 11;
  Matches    8;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          7 PRPTPPRPI 15
              |||||||:
Db         20 PRPTPPRPL 28


RESULT 3
Q8YTH2
ID   Q8YTH2      PRELIMINARY;      PRT;   305 AA.
AC   Q8YTH2;
DT   01-MAR-2002 (TrEMBLrel. 20, Created)
DT   01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Hypothetical protein Alr2745.
GN   ALR2745.
OS   Anabaena sp. (strain PCC 7120).
OC   Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.
OX   NCBI_TaxID=103690;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21595285; PubMed=11759840;
RA   Kaneko T., Nakamura Y., Wolk C.P., Kuritz T., Sasamoto S.,
RA   Watanabe A., Iriguchi M., Ishikawa A., Kawashima K., Kimura T.,
RA   Kishida Y., Kohara M., Matsumoto M., Matsuno A., Muraki A.,
RA   Nakazaki N., Shimpo S., Sugimoto M., Takazawa M., Yamada M.,
RA   Yasuda M., Tabata S.;
RT   "Complete genomic sequence of the filamentous nitrogen-fixing
RT   cyanobacterium Anabaena sp. strain PCC 7120.";
RL   DNA Res. 8:205-213(2001).
DR   EMBL; AP003590; BAB74444.1; -.
DR   PIR; AB2149; AB2149.
DR   GO; GO:0000270; P:peptidoglycan metabolism; IEA.
DR   InterPro; IPR002477; PG_binding.
DR   Pfam; PF01471; PG_binding_1; 2.
KW   Hypothetical protein; Complete proteome.
SQ   SEQUENCE   305 AA;  32953 MW;  954FF5E2BDC0AC83 CRC64;

  Query Match             58.1%;  Score 50;  DB 16;  Length 305;
  Best Local Similarity   100.0%;  Pred. No. 8.9;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          7 PRPTPPRP 14
              ||||||||
Db        187 PRPTPPRP 194


RESULT 4
O94113
```

```
ID   O94113      PRELIMINARY;      PRT;   671 AA.
AC   O94113;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Kexin (Fragment).
GN   KEXIN.
OS   Pneumocystis carinii.
OC   Eukaryota; Fungi; Ascomycota; Pneumocystidomycetes; Pneumocystidaceae;
OC   Pneumocystis.
OX   NCBI_TaxID=4754;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Russian D.A., Andrawis-Sorial V., Angus C.W., Kovacs J.A.;
RL   Submitted (DEC-1996) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: CONTAINS 1 HOMO B/P DOMAIN.
DR   EMBL; U82999; AAD00541.1; -.
DR   MEROPS; S08.011; -.
DR   GO; GO:0004289; F:subtilase activity; IEA.
DR   GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR   InterPro; IPR000209; Peptidase_S8.
DR   InterPro; IPR002884; Peptidase_S8B.
DR   Pfam; PF00082; Peptidase_S8; 1.
DR   Pfam; PF01483; P_proprotein; 1.
DR   PRINTS; PR00723; SUBTILISIN.
DR   ProDom; PD000717; P_domain; 1.
DR   PROSITE; PS00137; SUBTILASE_HIS; 1.
DR   PROSITE; PS00138; SUBTILASE_SER; 1.
FT   NON_TER       1      1
SQ   SEQUENCE   671 AA;  74049 MW;  BAC4A164EC007C2E CRC64;

  Query Match             58.1%;  Score 50;  DB 3;  Length 671;
  Best Local Similarity   100.0%;  Pred. No. 19;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          7 PRPTPPRP 14
              ||||||||
Db        508 PRPTPPRP 515


RESULT 5
Q7ZX30
ID   Q7ZX30      PRELIMINARY;      PRT;   377 AA.
AC   Q7ZX30;
DT   01-JUN-2003 (TrEMBLrel. 24, Created)
DT   01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Similar to splicing factor 3b, subunit 4, 49 kDa.
OS   Xenopus laevis (African clawed frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
OC   Xenopodinae; Xenopus.
OX   NCBI_TaxID=8355;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Embryo;
RA   Klein S., Strausberg R.;
RL   Submitted (JAN-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC045264; AAH45264.1; -.
DR   GO; GO:0003676; F:nucleic acid binding; IEA.
DR   InterPro; IPR000504; RNA_rec_mot.
DR   Pfam; PF00076; rrm; 2.
DR   SMART; SM00360; RRM; 2.
DR   PROSITE; PS50102; RRM; 2.
DR   PROSITE; PS00030; RRM_RNP_1; 1.
SQ   SEQUENCE   377 AA;  40209 MW;  879D30A7A5FE022F CRC64;

  Query Match             57.0%;  Score 49;  DB 13;  Length 377;
  Best Local Similarity   88.9%;  Pred. No. 15;
  Matches    8;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          6 LPRPTPPRP 14
              |||| ||||
Db        361 LPRPAPPRP 369


RESULT 6
Q8L462
ID   Q8L462      PRELIMINARY;      PRT;   532 AA.
AC   Q8L462;
DT   01-OCT-2002 (TrEMBLrel. 22, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Putative no apical meristem (NAM) protein.
GN   OSJNBA0011L09.4 OR OSJNAA0011L09.8.
OS   Oryza sativa (japonica cultivar-group).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC   Ehrhartoideae; Oryzeae; Oryza.
OX   NCBI_TaxID=39947;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Nipponbare;
RA   Buell C.R., Yuan Q., Ouyang S., Liu J., Gansberger K., Kim M.M.,
RA   Overton II L.L., Bera J.J., Tsitrin T., Krol M.I., Jarrahi B.B.,
RA   Jin S.S., Koo H., Zismann V., Hsiao J., Blunt S., Vanaken S.S.,
RA   Utterback T.T., Feldblyum T.V., Yang Q.Q., Haas B.J., Suh B.B.,
RA   Peterson J.J., Quackenbush J., White O., Salzberg S.L., Fraser C.M.;
RT   "Oryza sativa chromosome 10 BAC OSJNBa0011L09 genomic sequence.";
RL   Submitted (MAY-2002) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Wing R.A., Yu Y., Yang T.J., Nah G., Soderlund C., Chen M., Kim H.-R.,
RA   Rambo T., Saski C., Henry D., Oates R., Simmons J.;
RT   "Rice Genomic Sequence.";
RL   Submitted (MAY-2002) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Nipponbare;
RA   The Rice Chromosome 10 Sequencing Consortium;
RT   "In-depth view of structure, activity, and evolution of rice
RT   chromosome 10.";
RL   Science 300:1566-1569(2003).
RN   [4]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Nipponbare;
RA   Buell C.R., Wing R.A., McCombie W.R., Messing J., Yuan Q.;
RL   Submitted (MAY-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AC092388; AAM22718.1; -.
DR   EMBL; AC122144; AAM44885.1; -.
DR   EMBL; AE017090; AAP53600.1; -.
DR   Gramene; Q8L462; -.
DR   InterPro; IPR003441; NAM.
DR   Pfam; PF02365; NAM; 1.
SQ   SEQUENCE   532 AA;  57452 MW;  AFD5B73997E2F2C2 CRC64;

  Query Match             57.0%;  Score 49;  DB 10;  Length 532;
  Best Local Similarity   80.0%;  Pred. No. 22;
  Matches    8;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          6 LPRPTPPRPI 15
              || ||||||:
Db        220 LPPPTPPRPL 229


RESULT 7
Q9DCD5
ID   Q9DCD5      PRELIMINARY;      PRT;   539 AA.
AC   Q9DCD5;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   0610041D19Rik protein.
GN   0610041D19RIK.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```

```
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Kidney;
RX   MEDLINE=21085660; PubMed=11217851;
RA   Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA   Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA   Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA   Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA   Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA   Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA   Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA   Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA   Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA   Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA   Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA   Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA   Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA   Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA   Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA   Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA   Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA   Hayashizaki Y.;
RT   "Functional annotation of a full-length mouse cDNA collection.";
RL   Nature 409:685-690(2001).
DR   EMBL; AK002882; BAB22429.1; -.
DR   MGD; MGI:1921344; 0610041D19Rik.
SQ   SEQUENCE   539 AA;  59419 MW;  FD169E54AD29E4DA CRC64;

  Query Match             55.8%;  Score 48;  DB 11;  Length 539;
  Best Local Similarity   50.0%;  Pred. No. 31;
  Matches    8;  Conservative    2;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 DKGXXLPRPTPPRPIY 16
              ||    | |:|| |:|
Db        296 DKLSPYPTPSPPHPLY 311


RESULT 8
Q8CFL7
ID   Q8CFL7      PRELIMINARY;      PRT;   539 AA.
AC   Q8CFL7;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Similar to RIKEN cDNA 0610041D19 gene.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Breast tumor;
RA   Strausberg R.;
RL   Submitted (FEB-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC023316; AAH23316.1; -.
DR   PIR; PT0675; PT0675.
SQ   SEQUENCE   539 AA;  59435 MW;  F764942CC75BE4C8 CRC64;

  Query Match             55.8%;  Score 48;  DB 11;  Length 539;
  Best Local Similarity   50.0%;  Pred. No. 31;
  Matches    8;  Conservative    2;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 DKGXXLPRPTPPRPIY 16
              ||    | |:|| |:|
Db        296 DKLSPYPTPSPPHPLY 311


RESULT 9
Q8DK04
ID   Q8DK04      PRELIMINARY;      PRT;   957 AA.
AC   Q8DK04;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Translation initiation factor IF-2.
GN   INFB OR TLR1066.
OS   Synechococcus elongatus (Thermosynechococcus elongatus).
OC   Bacteria; Cyanobacteria; Chroococcales; Synechococcus.
OX   NCBI_TaxID=32046;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BP-1;
RX   MEDLINE=22225144; PubMed=12240834;
RA   Nakamura Y., Kaneko T., Sato S., Ikeuchi M., Katoh H., Sasamoto S.,
RA   Watanabe A., Iriguchi M., Kawashima K., Kimura T., Kishida Y.,
RA   Kiyokawa C., Kohara M., Matsumoto M., Matsuno A., Nakazaki N.,
RA   Shimpo S., Sugimoto M., Takeuchi C., Yamada M., Tabata S.;
RT   "Complete genome structure of the thermophilic cyanobacterium
RT   Thermosynechococcus elongatus BP-1.";
RL   DNA Res. 9:123-130(2002).
DR   EMBL; AP005372; BAC08619.1; -.
DR   GO; GO:0005525; F:GTP binding; IEA.
DR   GO; GO:0003746; F:translation elongation factor activity; IEA.
DR   GO; GO:0003743; F:translation initiation factor activity; IEA.
DR   GO; GO:0006414; P:translational elongation; IEA.
DR   GO; GO:0006413; P:translational initiation; IEA.
DR   InterPro; IPR004161; EFTU_D2.
DR   InterPro; IPR000795; EF_GTPbind.
DR   InterPro; IPR000178; IF2.
DR   InterPro; IPR006847; IF2_N.
DR   InterPro; IPR005225; Small_GTP.
DR   InterPro; IPR009000; Translat_factor.
DR   Pfam; PF00009; GTP_EFTU; 1.
DR   Pfam; PF03144; GTP_EFTU_D2; 2.
DR   Pfam; PF04760; IF2_N; 2.
DR   PRINTS; PR00315; ELONGATNFCT.
DR   ProDom; PD186100; IF2; 1.
DR   TIGRFAMs; TIGR00487; IF-2; 1.
DR   TIGRFAMs; TIGR00231; small_GTP; 1.
DR   PROSITE; PS01176; IF2; 1.
KW   Initiation factor; Complete proteome.
SQ   SEQUENCE   957 AA;  104247 MW;  13E9E041ADBC1280 CRC64;

  Query Match             55.8%;  Score 48;  DB 16;  Length 957;
  Best Local Similarity   77.8%;  Pred. No. 54;
  Matches    7;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          7 PRPTPPRPI 15
              | ||||||:
Db        144 PEPTPPRPV 152


RESULT 10
Q9NT81
ID   Q9NT81      PRELIMINARY;      PRT;  1013 AA.
AC   Q9NT81;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Hypothetical protein (Fragment).
GN   DKFZP434M2023.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Testis;
RA   Blum H., Bauersachs S., Mewes H.W., Gassenhuber J., Wiemann S.;
RL   Submitted (JAN-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AL137480; CAB70761.1; -.
DR   PIR; T46422; T46422.
DR   InterPro; IPR001202; WW_Rsp5_WWP.
DR   Pfam; PF00397; WW; 2.
```

```
DR   SMART; SM00456; WW; 2.
DR   PROSITE; PS01159; WW_DOMAIN_1; 1.
DR   PROSITE; PS50020; WW_DOMAIN_2; 2.
KW   Hypothetical protein.
FT   NON_TER       1      1
SQ   SEQUENCE   1013 AA;  109875 MW;  36F1E4507468C0FF CRC64;

  Query Match             55.8%;  Score 48;  DB 4;  Length 1013;
  Best Local Similarity   66.7%;  Pred. No. 58;
  Matches    8;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          3 GXXLPRPTPPRP 14
              |   | ||||||
Db        161 GASAPPPTPPRP 172


RESULT 11
Q8N3X1
ID   Q8N3X1      PRELIMINARY;      PRT;  1015 AA.
AC   Q8N3X1;
DT   01-OCT-2002 (TrEMBLrel. 22, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Similar to formin binding protein 4.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Cervix;
RA   Strausberg R.;
RL   Submitted (SEP-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC037404; AAH37404.1; -.
DR   InterPro; IPR001202; WW_Rsp5_WWP.
DR   Pfam; PF00397; WW; 2.
DR   SMART; SM00456; WW; 2.
DR   PROSITE; PS01159; WW_DOMAIN_1; 1.
DR   PROSITE; PS50020; WW_DOMAIN_2; 2.
SQ   SEQUENCE   1015 AA;  110063 MW;  3652A950746E3594 CRC64;

  Query Match             55.8%;  Score 48;  DB 4;  Length 1015;
  Best Local Similarity   66.7%;  Pred. No. 58;
  Matches    8;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          3 GXXLPRPTPPRP 14
              |   | ||||||
Db        163 GASAPPPTPPRP 174


RESULT 12
Q9Y2L7
ID   Q9Y2L7      PRELIMINARY;      PRT;  1050 AA.
AC   Q9Y2L7;
DT   01-NOV-1999 (TrEMBLrel. 12, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Hypothetical protein KIAA1014 (Fragment).
GN   KIAA1014.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Brain;
RX   MEDLINE=99246063; PubMed=10231032;
RA   Nagase T., Ishikawa K., Suyama M., Kikuno R., Hirosawa M.,
RA   Miyajima N., Tanaka A., Kotani H., Nomura N., Ohara O.;
RT   "Prediction of the coding sequences of unidentified human genes. XIII.
RT   The complete sequences of 100 new cDNA clones from brain which code
RT   for large proteins in vitro.";
RL   DNA Res. 6:63-70(1999).
DR   EMBL; AB023231; BAA76858.2; -.
DR   Genew; HGNC:19752; FNBP4.
DR   InterPro; IPR001202; WW_Rsp5_WWP.
DR   Pfam; PF00397; WW; 2.
DR   SMART; SM00456; WW; 2.
DR   PROSITE; PS01159; WW_DOMAIN_1; 1.
DR   PROSITE; PS50020; WW_DOMAIN_2; 2.
KW   Hypothetical protein.
FT   NON_TER       1      1
SQ   SEQUENCE   1050 AA;  113629 MW;  09DD1747406D89E4 CRC64;

  Query Match             55.8%;  Score 48;  DB 4;  Length 1050;
  Best Local Similarity   66.7%;  Pred. No. 60;
  Matches    8;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          3 GXXLPRPTPPRP 14
              |   | ||||||
Db        198 GASAPPPTPPRP 209


RESULT 13
Q9LVN1
ID   Q9LVN1      PRELIMINARY;      PRT;  1307 AA.
AC   Q9LVN1;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Gb|AAD23008.1.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Columbia;
RX   MEDLINE=20181125; PubMed=10718197;
RA   Sato S., Nakamura Y., Kaneko T., Katoh T., Asamizu E., Kotani H.,
RA   Tabata S.;
RT   "Structural analysis of Arabidopsis thaliana chromosome 5. X. Sequence
RT   features of the regions of 3,076,755 bp covered by sixty P1 and TAC
RT   clones.";
RL   DNA Res. 7:31-63(2000).
DR   EMBL; AB019228; BAA96907.1; -.
DR   GO; GO:0003779; F:actin binding; IEA.
DR   InterPro; IPR008973; C2_CaLB.
DR   InterPro; IPR003104; FH2.
DR   Pfam; PF02181; FH2; 1.
DR   SMART; SM00498; FH2; 1.
SQ   SEQUENCE   1307 AA;  144545 MW;  CFD60BFB9669FA2A CRC64;

  Query Match             55.8%;  Score 48;  DB 10;  Length 1307;
  Best Local Similarity   64.3%;  Pred. No. 74;
  Matches    9;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          1 DKGXXLPRPTPPRP 14
              ||   |||| || |
Db        681 DKKPALPRPPPPPP 694


RESULT 14
Q9X9V2
ID   Q9X9V2      PRELIMINARY;      PRT;   165 AA.
AC   Q9X9V2;
DT   01-NOV-1999 (TrEMBLrel. 12, Created)
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Putative mini-circle protein.
GN   SCO5093 OR SCBAC28G1.19C.
OS   Streptomyces coelicolor.
OC   Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC   Streptomycineae; Streptomycetaceae; Streptomyces.
OX   NCBI_TaxID=1902;
```

```
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=A3(2);
RX   MEDLINE=99328982; PubMed=10400594;
RA   Martinez-Costa O.H., Martin-Triana A.J., Martinez E.,
RA   Fernandez-Moreno M.A., Malpartida F.;
RT   "An additinal regulatory gene for actinorhodin production in
RT   Streptomyces lividans involves a LysR-type transcriptional
RT   regulator.";
RL   J. Bacteriol. 181:4353-4364(1999).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=A3(2);
RA   Warren T., Harris D.;
RL   Submitted (JUL-2001) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=A3(2);
RA   Cerdeno A.M., Parkhill J., Barrell B.G., Rajandream M.A.;
RL   Submitted (JUL-2001) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   SEQUENCE FROM N.A.
RC   STRAIN=A3(2);
RX   MEDLINE=97000351; PubMed=8843436;
RA   Redenbach M., Kieser H.M., Denapaite D., Eichner A., Cullum J.,
RA   Kinashi H., Hopwood D.A.;
RT   "A set of ordered cosmids and a detailed genetic and physical map for
RT   the 8 Mb Streptomyces coelicolor A3(2) chromosome.";
RL   Mol. Microbiol. 21:77-96(1996).
RN   [5]
RP   SEQUENCE FROM N.A.
RC   STRAIN=A3(2) / M145;
RX   MEDLINE=21996410; PubMed=12000953;
RA   Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,
RA   Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,
RA   Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
RA   Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
RA   Huang C.-H., Kieser T., Larke L., Murphy L., Oliver K., O'Neil S.,
RA   Rabbinowitsch E., Rajandream M.A., Rutherford K., Rutter S.,
RA   Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
RA   Warren T., Wietzorrek A., Woodward J., Barrell B.G., Parkhill J.,
RA   Hopwood D.A.;
RT   "Complete genome sequence of the model actinomycete Streptomyces
RT   coelicolor A3(2).";
RL   Nature 417:141-147(2002).
DR   EMBL; Y18817; CAB51132.1; -.
DR   EMBL; AL939122; CAC44206.1; -.
DR   PIR; T45271; T45271.
DR   InterPro; IPR007061; DUF664.
DR   Pfam; PF04978; DUF664; 1.
KW   Complete proteome.
SQ   SEQUENCE   165 AA;  17924 MW;  B18201686ACC2D89 CRC64;

  Query Match          54.7%;  Score 47;  DB 16;  Length 165;
  Best Local Similarity  64.3%;  Pred. No. 14;
  Matches    9;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          1 DKGXXLPRPTPPRP 14
              | |  ||||  | ||
Db        117 DLGAPLPRPRPGRP 130


RESULT 15
Q9VRU9
ID   Q9VRU9      PRELIMINARY;      PRT;   239 AA.
AC   Q9VRU9;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   CG12330 protein.
GN   CG12330.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Berkeley;
RX   MEDLINE=20196006; PubMed=10731132;
RA   Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA   Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA   George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA   Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA   Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA   Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA   Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA   Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA   Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA   Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA   Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA   Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA   de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA   Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA   Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA   Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA   Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA   Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA   Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA   Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA   Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA   Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA   Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA   Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA   Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA   Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA   Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA   Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA   Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA   Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA   Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA   Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA   Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA   Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA   Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA   Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT   "The genome sequence of Drosophila melanogaster.";
RL   Science 287:2185-2195(2000).
DR   EMBL; AE003563; AAF50683.1; -.
DR   FlyBase; FBgn0035686; CG12330.
DR   InterPro; IPR000618; Insect_cuticle.
DR   Pfam; PF00379; Chitin_bind_4; 1.
DR   PRINTS; PR00947; CUTICLE.
DR   PROSITE; PS00233; CUTICLE; 1.
SQ   SEQUENCE   239 AA;  24412 MW;  98FDF199821EACF9 CRC64;

  Query Match          54.7%;  Score 47;  DB 5;  Length 239;
  Best Local Similarity  53.3%;  Pred. No. 20;
  Matches    8;  Conservative    1;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          2 KGXXLPRPTPPRPIY 16
              ||  | | ||:| |
Db         59 KGNGKPPPAPPKPSY 73


Search completed: March 11, 2004, 16:55:22
Job time : 65 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: March 11, 2004, 17:17:09 ; Search time 55 Seconds
(without alignments)
92.470 Million cell updates/sec

Title: US-09-980-804-1
Perfect score: 86
Sequence: 1 DKGXXLPRPTPPRPIYXX 18

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	81	94.2	19	6	ABG73945	Abg73945 Cell wall
2	81	94.2	20	2	AAR50300	Aar50300 Anti-bact
3	81	94.2	20	4	AAG62734	Aag62734 Amino aci
4	81	94.2	20	4	AAY72457	Aay72457 Pyrrhocor
5	81	94.2	20	4	AAY72455	Aay72455 Pyrrhocor
6	81	94.2	20	4	AAY72442	Aay72442 Pyrrhocor
7	81	94.2	20	4	AAY72441	Aay72441 Pyrrhocor
8	81	94.2	20	4	AAY72443	Aay72443 Pyrrhocor
9	81	94.2	20	4	AAY72447	Aay72447 Pyrrhocor
10	81	94.2	20	4	AAY72453	Aay72453 Pyrrhocor
11	81	94.2	20	4	AAY72498	Aay72498 Pyrrhocor
12	81	94.2	20	4	AAY72456	Aay72456 Pyrrhocor
13	81	94.2	20	4	AAY72437	Aay72437 Pyrrhocor
14	81	94.2	20	4	AAY72433	Aay72433 Native Py
15	81	94.2	20	4	AAY72435	Aay72435 Pyrrhocor
16	81	94.2	20	8	ADD35367	Add35367 Antimicro
17	81	94.2	21	4	AAG62743	Aag62743 Amino aci
18	81	94.2	21	4	AAG62756	Aag62756 Amino aci
19	81	94.2	21	4	AAY72439	Aay72439 Pyrrhocor
20	81	94.2	21	4	AAY72444	Aay72444 Pyrrhocor
21	81	94.2	21	4	AAY72454	Aay72454 Pyrrhocor
22	81	94.2	21	4	AAY72448	Aay72448 Pyrrhocor
23	81	94.2	21	4	AAY72440	Aay72440 Pyrrhocor
24	81	94.2	21	4	AAY72451	Aay72451 Pyrrhocor
25	81	94.2	21	4	AAY72452	Aay72452 Pyrrhocor
26	81	94.2	21	4	AAY72450	Aay72450 Pyrrhocor
27	81	94.2	21	4	AAY72445	Aay72445 Pyrrhocor
28	81	94.2	21	4	AAY72446	Aay72446 Pyrrhocor
29	81	94.2	23	4	AAY72461	Aay72461 Pyrrhocor
30	81	94.2	24	4	AAY72438	Aay72438 Pyrrhocor
31	81	94.2	29	4	AAY72449	Aay72449 Pyrrhocor
32	80	93.0	18	4	AAG62740	Aag62740 Amino aci
33	80	93.0	18	4	AAY72424	Aay72424 Pyrrhocor
34	74	86.0	18	4	AAG62767	Aag62767 Amino aci
35	52	60.5	133	6	ABO00915	Abo00915 Polypepti
36	50	58.1	115	7	ADC33332	Adc33332 Human nov
37	48	55.8	530	4	AAB64386	Aab64386 Amino aci
38	48	55.8	1017	4	AAM40352	Aam40352 Human pol
39	47	54.7	141	4	AAO03095	Aao03095 Human pol
40	47	54.7	176	5	AAU98529	Aau98529 Bcl-2 rel
41	47	54.7	239	4	ABB63164	Abb63164 Drosophil
42	47	54.7	538	5	ABG30975	Abg30975 Human 3-b
43	47	54.7	811	7	ADC31170	Adc31170 Human nov
44	47	54.7	863	3	AAB42952	Aab42952 Human ORF
45	47	54.7	903	3	AAB42926	Aab42926 Human ORF

ALIGNMENTS

RESULT 1
ABG73945
ID ABG73945 standard; peptide; 19 AA.
XX
AC ABG73945;
XX
DT 31-MAR-2003 (first entry)
XX
DE Cell wall/cell membrane transport peptide #4.
XX
KW Transport peptide; cell wall; cell membrane; protein nucleic acid; PNA;
KW genetically modified micro-organism; bacterial infection.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Modified-site 1
FT /label= OTHER
FT /note= "Lys is hydrogenated"
FT Modified-site 19
FT /label= OTHER
FT /note= "Cys is covalently linked via an smcc (not
FT defined) polyethylene-glycol moiety to the nucleic acid
FT sequence appearing as ABX15985"
XX
PN WO200279467-A2.
XX
PD 10-OCT-2002.
XX
PF 26-MAR-2002; 2002WO-DK000208.
XX
PR 29-MAR-2001; 2001DK-00000523.
XX
PA (UYKO-) UNIV KOBENHAVNS.
XX
PI Nielsen PE, Good L;
XX
DR WPI; 2003-103273/09.
XX
PT Selecting genetically modified cells useful for isolation and industrial
PT growth of transformed organisms comprises treating the modified cells
PT with an antisense or antigene construct directed against the essential
PT gene X of the cells.
XX
PS Claim 16; Page 51; 92pp; English.
XX
CC The invention relates to selecting genetically modified cells comprising:
CC (a) modifying cells containing a growth essential gene X, with a vector

CC containing gene Y; and (b) treating the modified cells with an antisense
CC or antigene construct directed against the essential gene X of the cells
CC to obtain preferential growth of the modified cells over other non-
CC modified cells. Also included is a product manufactured fully or
CC partially by use of the new method. The method is useful for selecting
CC genetically modified cells and manufacturing a product. It is useful for
CC research the isolation and industrial growth maintenance of transformed
CC organisms. The new method has the advantage of selecting and maintaining
CC a plasmid containing bacterial culture without the use of antibiotics.
CC This has a wide variety of applications in research, development, and
CC industrial production involving genetically modified micro-organisms. The
CC method inhibits bacterial infections in eukaryotic cell cultures. The
CC present sequence is a cell wall/cell membrane transport peptide which is
CC incorporated into a peptide nucleic acid (PNA) antisense molecules for
CC use in the method of the invention
XX
SQ Sequence 19 AA;

Query Match 94.2%; Score 81; DB 6; Length 19;
Best Local Similarity 87.5%; Pred. No. 0.00047;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 DKGXXLPRPTPPRPIY 16
||| ||||||||||
Db 2 DKGSYLPRPTPPRPIY 17

RESULT 2
AAR50300
ID AAR50300 standard; peptide; 20 AA.
XX
AC AAR50300;
XX
DT 25-MAR-2003 (revised)
DT 10-OCT-1994 (first entry)
XX
DE Anti-bacterial glycopeptide #9 induced in Pyrrhocoris apterus.
XX
KW Antibacterial glycopeptide; Diptera; septicaemia; Gram positive bacteria;
KW Gram negative bacteria.
XX
OS Pyrrhocoris apterus.
XX
FH Key Location/Qualifiers
FT Modified-site 11
FT /label= O-glycosylated
XX
PN WO9405787-A1.
XX
PD 17-MAR-1994.
XX
PF 06-SEP-1993; 93WO-FR000853.
XX
PR 04-SEP-1992; 92FR-00010608.
XX
PA (CNRS) CNRS CENT NAT RECH SCI.
XX
PI Bulet P, Hetru C, Dimarcq J, Hoffmann J, Van Dorsselaer A;
XX
DR WPI; 1994-101192/12.
XX
PT New antibacterial glycopeptide(s) derived from insects - for control of
PT Gram negative and positive bacteria in human and veterinary medicine,
PT agriculture, etc.
XX
PS Claim 17; Page 9-10; 45pp; French.
XX
CC This is a preferred example of an anti-bacterial glycopeptide induced in
CC arthropods (esp. larval or adult insects) by injection of bacteria, a
CC septic wound or other injury. The peptides contain at least one O-
CC glycosylated residue and are useful for treatment of e.g. septicaemia,
CC for oral or dental use and in gynaecology. (Updated on 25-MAR-2003 to
CC correct PN field.)
XX
SQ Sequence 20 AA;

Query Match 94.2%; Score 81; DB 2; Length 20;
Best Local Similarity 87.5%; Pred. No. 0.0005;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 DKGXXLPRPTPPRPIY 16
||| ||||||||||
Db 2 DKGSYLPRPTPPRPIY 17

RESULT 3
AAG62734
ID AAG62734 standard; peptide; 20 AA.
XX
AC AAG62734;
XX
DT 17-SEP-2001 (first entry)
XX
DE Amino acid sequence of antibacterial peptide pyrrhocoricin.
XX
KW Multi-helical lid; heat shock protein; hsp; protein folding;
KW pathogenic infection; bacterial infection; antibacterial.
XX
OS Unidentified.
XX
PN WO200153509-A2.
XX
PD 26-JUL-2001.
XX
PF 19-JAN-2001; 2001WO-US001812.
XX
PR 21-JAN-2000; 2000US-0177565P.
PR 03-OCT-2000; 2000US-0237599P.
XX
PA (WIST-) WISTAR INST ANATOMY & BIOLOGY.
PA (UYCR-) UNIV CREIGHTON.
XX
PI Otvos L, Blaszczyk-Thurin M, Rogers M, Lovas S;
XX
DR WPI; 2001-451911/48.
XX
PT Composition, used to treat a pathogenic infection and eliminate a plant,
PT insect, or animal pest, comprises a molecule that binds to a heat shock
PT protein.
XX
PS Example 6; Page 64; 124pp; English.
XX
CC The specification describes a composition that comprises a synthetic non-
CC naturally occurring molecule that binds to a selected multi-helical lid
CC of a heat shock protein (hsp) of a selected organism, where the molecule
CC inhibits protein folding activity of the hsp, and a carrier, where
CC exposure of the organism to the composition retards the growth and
CC reproduction of the organism. The composition is used to treat a mammal
CC suffering from a pathogenic infection, in the manufacture of a medicament
CC for treating a mammal for a pathogenic infection, and to eliminate a
CC plant, insect, or animal pest. It is used in the manufacture of a
CC medicament for treating mammalian bacterial infection. The present
CC sequence represents an antibacterial peptide, which may be used to
CC produce the composition of the invention
XX
SQ Sequence 20 AA;

Query Match 94.2%; Score 81; DB 4; Length 20;
Best Local Similarity 87.5%; Pred. No. 0.0005;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 DKGXXLPRPTPPRPIY 16
||| ||||||||||
Db 2 DKGSYLPRPTPPRPIY 17

```
RESULT 4
AAY72457
ID   AAY72457 standard; peptide; 20 AA.
XX
AC   AAY72457;
XX
DT   06-AUG-2003  (revised)
DT   24-APR-2001  (first entry)
XX
DE   Pyrrhocoricin-modified Peptide 13.
XX
KW   Pyrrhocoricin-derived peptide; antibacterial; fungicidal; therapy;
KW   fungal infection; bacterial infection; candidiasis; drug development.
XX
OS   Pyrrhocoris apterus.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 1. .20
FT                   /note= "D-form residues"
XX
PN   WO200078956-A1.
XX
PD   28-DEC-2000.
XX
PF   21-JUN-2000; 2000WO-US016989.
XX
PR   23-JUN-1999;   99US-0140606P.
PR   15-SEP-1999;   99US-0154135P.
XX
PA   (WIST-) WISTAR INST ANATOMY & BIOLOGY.
XX
PI   Otvos L;
XX
DR   WPI; 2001-112323/12.
XX
PT   Polypeptides derived from the peptide pyrrhocoricin, useful for treating
PT   fungal infections and Gram negative/positive bacterial infections.
XX
PS   Example 1; Page 25; 75pp; English.
XX
CC   The present peptide sequence is inactive Pyrrhocoricin-modified Peptide
CC   13. Pyrrhocoricin is a glycopeptide characterised by the presence of a
CC   disaccharide in the mid-chain position. The invention relates to
CC   pyrrhocoricin-derived peptides which have anti-bacterial or anti-fungal
CC   activity. These peptides have metabolic stability in mammalian serum. The
CC   pyrrhocoricin-derived peptides are used in the treatment of bacterial
CC   infections caused by Gram positive or Gram negative bacterium and fungal
CC   infections of skin, nails, mucus membranes and intestines e.g.,
CC   candidiasis. These peptides are also useful in anti-bacterial or anti-
CC   fungal pharmaceutical compositions, drug development and identification
CC   of other antibiotic or anti-fungal compounds. (Updated on 06-AUG-2003 to
CC   correct OS field.)
XX
SQ   Sequence 20 AA;

  Query Match             94.2%;  Score 81;  DB 4;  Length 20;
  Best Local Similarity   87.5%;  Pred. No. 0.0005;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 DKGXXLPRPTPPRPIY 16
              |||  ||||||||||
Db          2 DKGSYLPRPTPPRPIY 17


RESULT 5
AAY72455
ID   AAY72455 standard; peptide; 20 AA.
XX
AC   AAY72455;
XX
DT   06-AUG-2003  (revised)
DT   24-APR-2001  (first entry)
XX
DE   Pyrrhocoricin-modified Peptide 23.
XX
KW   Pyrrhocoricin-derived peptide; antibacterial; fungicidal; therapy;
KW   fungal infection; bacterial infection; candidiasis; drug development.
XX
OS   Pyrrhocoris apterus.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   1
FT                   /note= "Homoproline or 1-aminocyclo-hexane carboxylic
FT                   acid"
FT   Misc-difference 5
FT                   /note= "Wild type Ser substituted with Ala"
FT   Misc-difference 6
FT                   /note= "Wild type Tyr substituted with Phe"
FT   Modified-site   20
FT                   /note= "Beta-acetyl-2,3-diamino propionic acid"
XX
PN   WO200078956-A1.
XX
PD   28-DEC-2000.
XX
PF   21-JUN-2000; 2000WO-US016989.
XX
PR   23-JUN-1999;   99US-0140606P.
PR   15-SEP-1999;   99US-0154135P.
XX
PA   (WIST-) WISTAR INST ANATOMY & BIOLOGY.
XX
PI   Otvos L;
XX
DR   WPI; 2001-112323/12.
XX
PT   Polypeptides derived from the peptide pyrrhocoricin, useful for treating
PT   fungal infections and Gram negative/positive bacterial infections.
XX
PS   Example 1; Page 28; 75pp; English.
XX
CC   The present peptide sequence is inactive Pyrrhocoricin-modified Peptide
CC   23. Pyrrhocoricin is a glycopeptide characterised by the presence of a
CC   disaccharide in the mid-chain position. The invention relates to
CC   pyrrhocoricin-derived peptides which have anti-bacterial or anti-fungal
CC   activity. These peptides have metabolic stability in mammalian serum. The
CC   pyrrhocoricin-derived peptides are used in the treatment of bacterial
CC   infections caused by Gram positive or Gram negative bacterium and fungal
CC   infections of skin, nails, mucus membranes and intestines e.g.,
CC   candidiasis. These peptides are also useful in anti-bacterial or anti-
CC   fungal pharmaceutical compositions, drug development and identification
CC   of other antibiotic or anti-fungal compounds. (Updated on 06-AUG-2003 to
CC   correct OS field.)
XX
SQ   Sequence 20 AA;

  Query Match             94.2%;  Score 81;  DB 4;  Length 20;
  Best Local Similarity   87.5%;  Pred. No. 0.0005;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 DKGXXLPRPTPPRPIY 16
              |||  ||||||||||
Db          2 DKGAFLPRPTPPRPIY 17


RESULT 6
AAY72442
ID   AAY72442 standard; peptide; 20 AA.
XX
AC   AAY72442;
XX
DT   06-AUG-2003  (revised)
DT   24-APR-2001  (first entry)
XX
```

```
DE   Pyrrhocoricin-modified Peptide 7.
XX
KW   Pyrrhocoricin-derived peptide; antibacterial; fungicidal; therapy;
KW   fungal infection; bacterial infection; candidiasis; drug development.
XX
OS   Pyrrhocoris apterus.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   1
FT                   /note= "N-terminal acetyl"
FT   Modified-site   11
FT                   /note= "Modified with galactose-2-acetamido-2-deoxy-
FT                   galactose (Gal-GalNAc)"
XX
PN   WO200078956-A1.
XX
PD   28-DEC-2000.
XX
PF   21-JUN-2000; 2000WO-US016989.
XX
PR   23-JUN-1999;   99US-0140606P.
PR   15-SEP-1999;   99US-0154135P.
XX
PA   (WIST-) WISTAR INST ANATOMY & BIOLOGY.
XX
PI   Otvos L;
XX
DR   WPI; 2001-112323/12.
XX
PT   Polypeptides derived from the peptide pyrrhocoricin, useful for treating
PT   fungal infections and Gram negative/positive bacterial infections.
XX
PS   Example 1; Page 24; 75pp; English.
XX
CC   The present peptide sequence is inactive Pyrrhocoricin-modified Peptide
CC   7. Pyrrhocoricin is a glycopeptide characterised by the presence of a
CC   disaccharide in the mid-chain position. The invention relates to
CC   pyrrhocoricin-derived peptides which have anti-bacterial or anti-fungal
CC   activity. These peptides have metabolic stability in mammalian serum. The
CC   pyrrhocoricin-derived peptides are used in the treatment of bacterial
CC   infections caused by Gram positive or Gram negative bacterium and fungal
CC   infections of skin, nails, mucus membranes and intestines e.g.,
CC   candidiasis. These peptides are also useful in anti-bacterial or anti-
CC   fungal pharmaceutical compositions, drug development and identification
CC   of other antibiotic or anti-fungal compounds. (Updated on 06-AUG-2003 to
CC   correct OS field.)
XX
SQ   Sequence 20 AA;

  Query Match             94.2%;  Score 81;  DB 4;  Length 20;
  Best Local Similarity   87.5%;  Pred. No. 0.0005;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 DKGXXLPRPTPPRPIY 16
              |||  ||||||||||
Db          2 DKGSYLPRPTPPRPIY 17


RESULT 7
AAY72441
ID   AAY72441 standard; peptide; 20 AA.
XX
AC   AAY72441;
XX
DT   06-AUG-2003  (revised)
DT   24-APR-2001  (first entry)
XX
DE   Pyrrhocoricin-modified Peptide 6.
XX
KW   Pyrrhocoricin-derived peptide; antibacterial; fungicidal; therapy;
KW   fungal infection; bacterial infection; candidiasis; drug development.
XX
OS   Pyrrhocoris apterus.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   1
FT                   /note= "Homoproline or 1-aminocyclo-hexane carboxylic
FT                   acid"
XX
PN   WO200078956-A1.
XX
PD   28-DEC-2000.
XX
PF   21-JUN-2000; 2000WO-US016989.
XX
PR   23-JUN-1999;   99US-0140606P.
PR   15-SEP-1999;   99US-0154135P.
XX
PA   (WIST-) WISTAR INST ANATOMY & BIOLOGY.
XX
PI   Otvos L;
XX
DR   WPI; 2001-112323/12.
XX
PT   Polypeptides derived from the peptide pyrrhocoricin, useful for treating
PT   fungal infections and Gram negative/positive bacterial infections.
XX
PS   Claim 25; Page 45; 75pp; English.
XX
CC   The present peptide sequence is active Pyrrhocoricin-modified Peptide 6.
CC   Pyrrhocoricin is a glycopeptide characterised by the presence of a
CC   disaccharide in the mid-chain position. The invention relates to
CC   pyrrhocoricin-derived peptides which have anti-bacterial or anti-fungal
CC   activity. These peptides have metabolic stability in mammalian serum. The
CC   pyrrhocoricin-derived peptides are used in the treatment of bacterial
CC   infections caused by Gram positive or Gram negative bacterium and fungal
CC   infections of skin, nails, mucus membranes and intestines e.g.,
CC   candidiasis. These peptides are also useful in anti-bacterial or anti-
CC   fungal pharmaceutical compositions, drug development and identification
CC   of other antibiotic or anti-fungal compounds. (Updated on 06-AUG-2003 to
CC   correct OS field.)
XX
SQ   Sequence 20 AA;

  Query Match             94.2%;  Score 81;  DB 4;  Length 20;
  Best Local Similarity   87.5%;  Pred. No. 0.0005;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 DKGXXLPRPTPPRPIY 16
              |||  ||||||||||
Db          2 DKGSYLPRPTPPRPIY 17


RESULT 8
AAY72443
ID   AAY72443 standard; peptide; 20 AA.
XX
AC   AAY72443;
XX
DT   06-AUG-2003  (revised)
DT   24-APR-2001  (first entry)
XX
DE   Pyrrhocoricin-modified Peptide 8.
XX
KW   Pyrrhocoricin-derived peptide; antibacterial; fungicidal; therapy;
KW   fungal infection; bacterial infection; candidiasis; drug development.
XX
OS   Pyrrhocoris apterus.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   1
FT                   /note= "N-terminal acetyl"
FT   Modified-site   20
```

```
FT                  /note= "C-terminal imide"
XX
PN   WO200078956-A1.
XX
PD   28-DEC-2000.
XX
PF   21-JUN-2000; 2000WO-US016989.
XX
PR   23-JUN-1999;   99US-0140606P.
PR   15-SEP-1999;   99US-0154135P.
XX
PA   (WIST-) WISTAR INST ANATOMY & BIOLOGY.
XX
PI   Otvos L;
XX
DR   WPI; 2001-112323/12.
XX
PT   Polypeptides derived from the peptide pyrrhocoricin, useful for treating
PT   fungal infections and Gram negative/positive bacterial infections.
XX
PS   Claim 26; Page 45; 75pp; English.
XX
CC   The present peptide sequence is active Pyrrhocoricin-modified Peptide 8.
CC   Pyrrhocoricin is a glycopeptide characterised by the presence of a
CC   disaccharide in the mid-chain position. The invention relates to
CC   pyrrhocoricin-derived peptides which have anti-bacterial or anti-fungal
CC   activity. These peptides have metabolic stability in mammalian serum. The
CC   pyrrhocoricin-derived peptides are used in the treatment of bacterial
CC   infections caused by Gram positive or Gram negative bacterium and fungal
CC   infections of skin, nails, mucus membranes and intestines e.g.,
CC   candidiasis. These peptides are also useful in anti-bacterial or anti-
CC   fungal pharmaceutical compositions, drug development and identification
CC   of other antibiotic or anti-fungal compounds. (Updated on 06-AUG-2003 to
CC   correct OS field.)
XX
SQ   Sequence 20 AA;

  Query Match             94.2%;  Score 81;  DB 4;  Length 20;
  Best Local Similarity   87.5%;  Pred. No. 0.0005;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 DKGXXLPRPTPPRPIY 16
              |||  ||||||||||||
Db          3 DKGSYLPRPTPPRPIY 18


RESULT 9
AAY72447
ID   AAY72447 standard; peptide; 20 AA.
XX
AC   AAY72447;
XX
DT   06-AUG-2003  (revised)
DT   24-APR-2001  (first entry)
XX
DE   Pyrrhocoricin-modified Peptide 12.
XX
KW   Pyrrhocoricin-derived peptide; antibacterial; fungicidal; therapy;
KW   fungal infection; bacterial infection; candidiasis; drug development.
XX
OS   Pyrrhocoris apterus.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 1
FT                   /note= "D-form residue"
FT   Misc-difference 20
FT                   /note= "D-form residue"
XX
PN   WO200078956-A1.
XX
PD   28-DEC-2000.
XX
PF   21-JUN-2000; 2000WO-US016989.
XX
PR   23-JUN-1999;   99US-0140606P.
PR   15-SEP-1999;   99US-0154135P.
XX
PA   (WIST-) WISTAR INST ANATOMY & BIOLOGY.
XX
PI   Otvos L;
XX
DR   WPI; 2001-112323/12.
XX
PT   Polypeptides derived from the peptide pyrrhocoricin, useful for treating
PT   fungal infections and Gram negative/positive bacterial infections.
XX
PS   Claim 30; Page 46; 75pp; English.
XX
CC   The present peptide sequence is active Pyrrhocoricin-modified Peptide 12.
CC   Pyrrhocoricin is a glycopeptide characterised by the presence of a
CC   disaccharide in the mid-chain position. The invention relates to
CC   pyrrhocoricin-derived peptides which have anti-bacterial or anti-fungal
CC   activity. These peptides have metabolic stability in mammalian serum. The
CC   pyrrhocoricin-derived peptides are used in the treatment of bacterial
CC   infections caused by Gram positive or Gram negative bacterium and fungal
CC   infections of skin, nails, mucus membranes and intestines e.g.,
CC   candidiasis. These peptides are also useful in anti-bacterial or anti-
CC   fungal pharmaceutical compositions, drug development and identification
CC   of other antibiotic or anti-fungal compounds. (Updated on 06-AUG-2003 to
CC   correct OS field.)
XX
SQ   Sequence 20 AA;

  Query Match             94.2%;  Score 81;  DB 4;  Length 20;
  Best Local Similarity   87.5%;  Pred. No. 0.0005;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 DKGXXLPRPTPPRPIY 16
              |||  ||||||||||||
Db          2 DKGSYLPRPTPPRPIY 17


RESULT 10
AAY72453
ID   AAY72453 standard; peptide; 20 AA.
XX
AC   AAY72453;
XX
DT   06-AUG-2003  (revised)
DT   24-APR-2001  (first entry)
XX
DE   Pyrrhocoricin-modified Peptide 21.
XX
KW   Pyrrhocoricin-derived peptide; antibacterial; fungicidal; therapy;
KW   fungal infection; bacterial infection; candidiasis; drug development.
XX
OS   Pyrrhocoris apterus.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   1
FT                   /note= "Homoproline or 1-aminocyclo-hexane carboxylic
FT                   acid"
FT   Modified-site   20
FT                   /note= "Beta-acetyl-2,3-diamino propionic acid"
XX
PN   WO200078956-A1.
XX
PD   28-DEC-2000.
XX
PF   21-JUN-2000; 2000WO-US016989.
XX
PR   23-JUN-1999;   99US-0140606P.
PR   15-SEP-1999;   99US-0154135P.
XX
```

PA (WIST-) WISTAR INST ANATOMY & BIOLOGY.
XX
PI Otvos L;
XX
DR WPI; 2001-112323/12.
XX
PT Polypeptides derived from the peptide pyrrhocoricin, useful for treating
PT fungal infections and Gram negative/positive bacterial infections.
XX
PS Claim 34; Page 47; 75pp; English.
XX
CC The present peptide sequence is active Pyrrhocoricin-modified Peptide 21.
CC Pyrrhocoricin is a glycopeptide characterised by the presence of a
CC disaccharide in the mid-chain position. The invention relates to
CC pyrrhocoricin-derived peptides which have anti-bacterial or anti-fungal
CC activity. These peptides have metabolic stability in mammalian serum. The
CC pyrrhocoricin-derived peptides are used in the treatment of bacterial
CC infections caused by Gram positive or Gram negative bacterium and fungal
CC infections of skin, nails, mucus membranes and intestines e.g.,
CC candidiasis. These peptides are also useful in anti-bacterial or anti-
CC fungal pharmaceutical compositions, drug development and identification
CC of other antibiotic or anti-fungal compounds. (Updated on 06-AUG-2003 to
CC correct OS field.)
XX
SQ Sequence 20 AA;

Query Match 94.2%; Score 81; DB 4; Length 20;
Best Local Similarity 87.5%; Pred. No. 0.0005;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 DKGXXLPRPTPPRPIY 16
||| |||||||||||
Db 2 DKGSYLPRPTPPRPIY 17

RESULT 11
AAY72498
ID AAY72498 standard; peptide; 20 AA.
XX
AC AAY72498;
XX
DT 06-AUG-2003 (revised)
DT 24-APR-2001 (first entry)
XX
DE Pyrrhocoricin-modified peptide #2 for multi-peptide construction.
XX
KW Pyrrhocoricin-derived peptide; antibacterial; fungicidal; therapy;
KW fungal infection; bacterial infection; candidiasis; drug development.
XX
OS Pyrrhocoris apterus.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Modified-site 1
FT /note= "Homoproline or 1-aminocyclo-hexane carboxylic
FT acid"
FT Cross-links 20
FT /note= "The carboxy group of the 2-amino-3-acetylamino-
FT propionic acid residue 20 of AAY72498 is condensed onto
FT the side chain amino group of 2,3-diamino propionic acid
FT residue 20 of AAY72435 to cross link the two peptides
FT into a multipeptide"
FT Modified-site 20
FT /note= "2-amino-3-acetylamino-propionic acid residue"
XX
PN WO200078956-A1.
XX
PD 28-DEC-2000.
XX
PF 21-JUN-2000; 2000WO-US016989.
XX
PR 23-JUN-1999; 99US-0140606P.
PR 15-SEP-1999; 99US-0154135P.
XX
PA (WIST-) WISTAR INST ANATOMY & BIOLOGY.
XX
PI Otvos L;
XX
DR WPI; 2001-112323/12.
XX
PT Polypeptides derived from the peptide pyrrhocoricin, useful for treating
PT fungal infections and Gram negative/positive bacterial infections.
XX
PS Claim 51; Page 50; 75pp; English.
XX
CC The present peptide sequence is Pyrrhocoricin-modified peptide used for
CC multipeptide construction. Pyrrhocoricin is a glycopeptide characterised
CC by the presence of a disaccharide in the mid-chain position. The
CC invention relates to pyrrhocoricin-derived peptides which have anti-
CC bacterial or anti-fungal activity. These peptides have metabolic
CC stability in mammalian serum. The pyrrhocoricin-derived peptides are used
CC in the treatment of bacterial infections caused by Gram positive or Gram
CC negative bacterium and fungal infections of skin, nails, mucus membranes
CC and intestines e.g., candidiasis. These peptides are also useful in anti-
CC bacterial or anti-fungal pharmaceutical compositions, drug development
CC and identification of other antibiotic or anti-fungal compounds. (Updated
CC on 06-AUG-2003 to correct OS field.)
XX
SQ Sequence 20 AA;

Query Match 94.2%; Score 81; DB 4; Length 20;
Best Local Similarity 87.5%; Pred. No. 0.0005;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 DKGXXLPRPTPPRPIY 16
||| |||||||||||
Db 2 DKGSYLPRPTPPRPIY 17

RESULT 12
AAY72456
ID AAY72456 standard; peptide; 20 AA.
XX
AC AAY72456;
XX
DT 06-AUG-2003 (revised)
DT 24-APR-2001 (first entry)
XX
DE Pyrrhocoricin-modified Peptide 24.
XX
KW Pyrrhocoricin-derived peptide; antibacterial; fungicidal; therapy;
KW fungal infection; bacterial infection; candidiasis; drug development.
XX
OS Pyrrhocoris apterus.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Modified-site 20
FT /note= "Beta-acetyl-2,3-diamino propionic acid"
XX
PN WO200078956-A1.
XX
PD 28-DEC-2000.
XX
PF 21-JUN-2000; 2000WO-US016989.
XX
PR 23-JUN-1999; 99US-0140606P.
PR 15-SEP-1999; 99US-0154135P.
XX
PA (WIST-) WISTAR INST ANATOMY & BIOLOGY.
XX
PI Otvos L;
XX
DR WPI; 2001-112323/12.
XX
PT Polypeptides derived from the peptide pyrrhocoricin, useful for treating

PT fungal infections and Gram negative/positive bacterial infections.
XX
PS Claim 36; Page 47; 75pp; English.
XX
CC The present peptide sequence is active Pyrrhocoricin-modified Peptide 24.
CC Pyrrhocoricin is a glycopeptide characterised by the presence of a
CC disaccharide in the mid-chain position. The invention relates to
CC pyrrhocoricin-derived peptides which have anti-bacterial or anti-fungal
CC activity. These peptides have metabolic stability in mammalian serum. The
CC pyrrhocoricin-derived peptides are used in the treatment of bacterial
CC infections caused by Gram positive or Gram negative bacterium and fungal
CC infections of skin, nails, mucus membranes and intestines e.g.,
CC candidiasis. These peptides are also useful in anti-bacterial or anti-
CC fungal pharmaceutical compositions, drug development and identification
CC of other antibiotic or anti-fungal compounds. (Updated on 06-AUG-2003 to
CC correct OS field.)
XX
SQ Sequence 20 AA;

Query Match 94.2%; Score 81; DB 4; Length 20;
Best Local Similarity 87.5%; Pred. No. 0.0005;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 DKGXXLPRPTPPRPIY 16
||| |||||||||||
Db 2 DKGSYLPRPTPPRPIY 17

RESULT 13
AAY72437
ID AAY72437 standard; peptide; 20 AA.
XX
AC AAY72437;
XX
DT 06-AUG-2003 (revised)
DT 24-APR-2001 (first entry)
XX
DE Pyrrhocoricin-modified Peptide 1.
XX
KW Pyrrhocoricin-derived peptide; antibacterial; fungicidal; therapy;
KW fungal infection; bacterial infection; candidiasis; drug development.
XX
OS Pyrrhocoris apterus.
OS Synthetic.
XX
PN WO200078956-A1.
XX
PD 28-DEC-2000.
XX
PF 21-JUN-2000; 2000WO-US016989.
XX
PR 23-JUN-1999; 99US-0140606P.
PR 15-SEP-1999; 99US-0154135P.
XX
PA (WIST-) WISTAR INST ANATOMY & BIOLOGY.
XX
PI Otvos L;
XX
DR WPI; 2001-112323/12.
XX
PT Polypeptides derived from the peptide pyrrhocoricin, useful for treating
PT fungal infections and Gram negative/positive bacterial infections.
XX
PS Example 1; Page 23; 75pp; English.
XX
CC The present peptide sequence is active Pyrrhocoricin-modified Peptide 1
CC in which the naturally occurring mid-chain glycosylation is deleted.
CC Pyrrhocoricin is a glycopeptide characterised by the presence of a
CC disaccharide in the mid-chain position. The invention relates to
CC pyrrhocoricin-derived peptides which have anti-bacterial or anti-fungal
CC activity. These peptides have metabolic stability in mammalian serum. The
CC pyrrhocoricin-derived peptides are used in the treatment of bacterial
CC infections caused by Gram positive or Gram negative bacterium and fungal
CC infections of skin, nails, mucus membranes and intestines e.g.,
CC candidiasis. These peptides are also useful in anti-bacterial or anti-
CC fungal pharmaceutical compositions, drug development and identification
CC of other antibiotic or anti-fungal compounds. (Updated on 06-AUG-2003 to
CC correct OS field.)
XX
SQ Sequence 20 AA;

Query Match 94.2%; Score 81; DB 4; Length 20;
Best Local Similarity 87.5%; Pred. No. 0.0005;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 DKGXXLPRPTPPRPIY 16
||| |||||||||||
Db 2 DKGSYLPRPTPPRPIY 17

RESULT 14
AAY72433
ID AAY72433 standard; peptide; 20 AA.
XX
AC AAY72433;
XX
DT 06-AUG-2003 (revised)
DT 24-APR-2001 (first entry)
XX
DE Native Pyrrhocoricin, Peptide 2.
XX
KW Pyrrhocoricin-derived peptide; antibacterial; fungicidal; therapy;
KW fungal infection; bacterial infection; candidiasis; drug development.
XX
OS Pyrrhocoris apterus.
XX
FH Key Location/Qualifiers
FT Cleavage-site 5. .6
FT /label= Endopeptidase_cleavage_site
FT Modified-site 11
FT /note= "Modified with Galactose-2-acetamido-2- deoxy-
FT galactose (Gal-GalNAc)"
FT Cleavage-site 18. .19
FT /label= Endopeptidase_cleavage_site
XX
PN WO200078956-A1.
XX
PD 28-DEC-2000.
XX
PF 21-JUN-2000; 2000WO-US016989.
XX
PR 23-JUN-1999; 99US-0140606P.
PR 15-SEP-1999; 99US-0154135P.
XX
PA (WIST-) WISTAR INST ANATOMY & BIOLOGY.
XX
PI Otvos L;
XX
DR WPI; 2001-112323/12.
XX
PT Polypeptides derived from the peptide pyrrhocoricin, useful for treating
PT fungal infections and Gram negative/positive bacterial infections.
XX
PS Example 1; Page 23; 75pp; English.
XX
CC The present sequence is native pyrrhocoricin, Peptide 2 which is
CC glycosylated. Pyrrhocoricin is a glycopeptide characterised by the
CC presence of a disaccharide in the mid-chain position. The invention
CC relates to pyrrhocoricin-derived peptides which have anti-bacterial or
CC anti-fungal activity. These peptides have metabolic stability in
CC mammalian serum. The pyrrhocoricin-derived peptides are used in the
CC treatment of bacterial infections caused by Gram positive or Gram
CC negative bacterium and fungal infections of skin, nails, mucus membranes
CC and intestines e.g., candidiasis. These peptides are also useful in anti-
CC bacterial or anti-fungal pharmaceutical compositions, drug development
CC and identification of other antibiotic or anti-fungal compounds. (Updated

CC on 06-AUG-2003 to correct OS field.)
XX
SQ Sequence 20 AA;

Query Match 94.2%; Score 81; DB 4; Length 20;
Best Local Similarity 87.5%; Pred. No. 0.0005;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 DKGXXLPRPTPPRPIY 16
||| ||||||||||
Db 2 DKGSYLPRPTPPRPIY 17

RESULT 15
AAY72435
ID AAY72435 standard; peptide; 20 AA.
XX
AC AAY72435;
XX
DT 06-AUG-2003 (revised)
DT 24-APR-2001 (first entry)
XX
DE Pyrrhocoricin-modified peptide #1 for multi-peptide construction.
XX
KW Pyrrhocoricin-derived peptide; antibacterial; fungicidal; therapy;
KW fungal infection; bacterial infection; candidiasis; drug development.
XX
OS Pyrrhocoris apterus.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Modified-site 1
FT /note= "Homoproline or 1-aminocyclo-hexane carboxylic
FT acid"
FT Cross-links 20
FT /note= "The carboxy group of the 2-amino-3-acetylamino-
FT propionic acid residue 20 of AAY72498 is condensed onto
FT the side chain amino group of 2,3-diamino propionic acid
FT residue 20 of AAY72435 to cross link the two peptides
FT into a multipeptide"
FT Modified-site 20
FT /note= "2,3-diamino propionic acid amide"
XX
PN WO200078956-A1.
XX
PD 28-DEC-2000.
XX
PF 21-JUN-2000; 2000WO-US016989.
XX
PR 23-JUN-1999; 99US-0140606P.
PR 15-SEP-1999; 99US-0154135P.
XX
PA (WIST-) WISTAR INST ANATOMY & BIOLOGY.
XX
PI Otvos L;
XX
DR WPI; 2001-112323/12.
XX
PT Polypeptides derived from the peptide pyrrhocoricin, useful for treating
PT fungal infections and Gram negative/positive bacterial infections.
XX
PS Claim 51; Page 50; 75pp; English.
XX
CC The present peptide sequence is Pyrrhocoricin-modified peptide used for
CC multipeptide construction. Pyrrhocoricin is a glycopeptide characterised
CC by the presence of a disaccharide in the mid-chain position. The
CC invention relates to pyrrhocoricin-derived peptides which have anti-
CC bacterial or anti-fungal activity. These peptides have metabolic
CC stability in mammalian serum. The pyrrhocoricin-derived peptides are used
CC in the treatment of bacterial infections caused by Gram positive or Gram
CC negative bacterium and fungal infections of skin, nails, mucus membranes
CC and intestines e.g., candidiasis. These peptides are also useful in anti-
CC bacterial or anti-fungal pharmaceutical compositions, drug development
CC and identification of other antibiotic or anti-fungal compounds. (Updated
CC on 06-AUG-2003 to correct OS field.)
XX
SQ Sequence 20 AA;

Query Match 94.2%; Score 81; DB 4; Length 20;
Best Local Similarity 87.5%; Pred. No. 0.0005;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 DKGXXLPRPTPPRPIY 16
||| ||||||||||
Db 2 DKGSYLPRPTPPRPIY 17

RESULT 16
ADD35367
ID ADD35367 standard; peptide; 20 AA.
XX
AC ADD35367;
XX
DT 15-JAN-2004 (first entry)
XX
DE Antimicrobial peptide pyrrhocoricin.
XX
KW antimicrobial; ophthalmic; prostaglandin; hypotensive; ophthalmological;
KW intraocular pressure; glaucoma; ocular hypertension; hyperaemia;
KW irritation; inflammation; conjunctiva; ocular cell dysplasia;
KW iridial melanocyte hyperplasia; hyperpigmentation.
XX
OS Unidentified.
XX
PN WO2003079997-A2.
XX
PD 02-OCT-2003.
XX
PF 21-MAR-2003; 2003WO-US008935.
XX
PR 21-MAR-2002; 2002US-0367071P.
XX
PA (CAYM-) CAYMAN CHEM CO.
XX
PI Maxey KM, Johnson J;
XX
DR WPI; 2004-011506/01.
XX
PT Ophthalmic solution useful for the treatment of increased intraocular
PT pressure comprises a prostaglandin of the F-series and an antimicrobial
PT peptide.
XX
PS Disclosure; Page 11; 11pp; English.
XX
CC The invention relates to a novel ophthalmic solution comprising a
CC prostaglandin of the F-series and an antimicrobial peptide. A solution of
CC the invention has hypotensive and ophthalmological activity. The solution
CC is useful for the treatment of increased intraocular pressure, such as
CC caused by glaucoma and for the reduction of ocular hypertension. The
CC prostaglandin and the antimicrobial peptide work synergistically, to
CC provide beneficial reduction in the incidence of irritant and toxic side
CC effects such as hyperaemia, irritation and inflammation of conjunctiva,
CC ocular cell dysplasia, iridial melanocyte hyperplasia, and
CC hyperpigmentation, associated with the prior art prostaglandin
CC compositions. The present sequence represents an antimicrobial peptide of
CC the invention.
XX
SQ Sequence 20 AA;

Query Match 94.2%; Score 81; DB 8; Length 20;
Best Local Similarity 87.5%; Pred. No. 0.0005;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 DKGXXLPRPTPPRPIY 16
||| ||||||||||
Db 2 DKGSYLPRPTPPRPIY 17

```
RESULT 17
AAG62743
ID   AAG62743 standard; peptide; 21 AA.
XX
AC   AAG62743;
XX
DT   17-SEP-2001  (first entry)
XX
DE   Amino acid sequence of modified antibacterial peptide pyrrhocoricin.
XX
KW   Multi-helical lid; heat shock protein; hsp; protein folding;
KW   pathogenic infection; bacterial infection; antibacterial.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   1
FT                   /note= "biotin attached"
XX
PN   WO200153509-A2.
XX
PD   26-JUL-2001.
XX
PF   19-JAN-2001; 2001WO-US001812.
XX
PR   21-JAN-2000; 2000US-0177565P.
PR   03-OCT-2000; 2000US-0237599P.
XX
PA   (WIST-) WISTAR INST ANATOMY & BIOLOGY.
PA   (UYCR-) UNIV CREIGHTON.
XX
PI   Otvos L,  Blaszczyk-Thurin M,  Rogers M,  Lovas S;
XX
DR   WPI; 2001-451911/48.
XX
PT   Composition, used to treat a pathogenic infection and eliminate a plant,
PT   insect, or animal pest, comprises a molecule that binds to a heat shock
PT   protein.
XX
PS   Example 1; Page 46; 124pp; English.
XX
CC   The specification describes a composition that comprises a synthetic non-
CC   naturally occurring molecule that binds to a selected multi-helical lid
CC   of a heat shock protein (hsp) of a selected organism, where the molecule
CC   inhibits protein folding activity of the hsp, and a carrier, where
CC   exposure of the organism to the composition retards the growth and
CC   reproduction of the organism. The composition is used to treat a mammal
CC   suffering from a pathogenic infection, in the manufacture of a medicament
CC   for treating a mammal for a pathogenic infection, and to eliminate a
CC   plant, insect, or animal pest. It is used in the manufacture of a
CC   medicament for treating mammalian bacterial infection. The present
CC   sequence represents a modified antibacterial peptide, which may be used
CC   to produce the composition of the invention
XX
SQ   Sequence 21 AA;

  Query Match             94.2%;  Score 81;  DB 4;  Length 21;
  Best Local Similarity   87.5%;  Pred. No. 0.00052;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 DKGXXLPRPTPPRPIY 16
              |||  ||||||||||||
Db          3 DKGSYLPRPTPPRPIY 18


RESULT 18
AAG62756
ID   AAG62756 standard; peptide; 21 AA.
XX
AC   AAG62756;
XX
DT   17-SEP-2001  (first entry)
XX
DE   Amino acid sequence of modified antibacterial peptide pyrrhocoricin.
XX
KW   Multi-helical lid; heat shock protein; hsp; protein folding;
KW   pathogenic infection; bacterial infection; antibacterial.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   1
FT                   /note= "fluorescein attached"
XX
PN   WO200153509-A2.
XX
PD   26-JUL-2001.
XX
PF   19-JAN-2001; 2001WO-US001812.
XX
PR   21-JAN-2000; 2000US-0177565P.
PR   03-OCT-2000; 2000US-0237599P.
XX
PA   (WIST-) WISTAR INST ANATOMY & BIOLOGY.
PA   (UYCR-) UNIV CREIGHTON.
XX
PI   Otvos L,  Blaszczyk-Thurin M,  Rogers M,  Lovas S;
XX
DR   WPI; 2001-451911/48.
XX
PT   Composition, used to treat a pathogenic infection and eliminate a plant,
PT   insect, or animal pest, comprises a molecule that binds to a heat shock
PT   protein.
XX
PS   Example 1; Page 56; 124pp; English.
XX
CC   The specification describes a composition that comprises a synthetic non-
CC   naturally occurring molecule that binds to a selected multi-helical lid
CC   of a heat shock protein (hsp) of a selected organism, where the molecule
CC   inhibits protein folding activity of the hsp, and a carrier, where
CC   exposure of the organism to the composition retards the growth and
CC   reproduction of the organism. The composition is used to treat a mammal
CC   suffering from a pathogenic infection, in the manufacture of a medicament
CC   for treating a mammal for a pathogenic infection, and to eliminate a
CC   plant, insect, or animal pest. It is used in the manufacture of a
CC   medicament for treating mammalian bacterial infection. The present
CC   sequence represents a modified antibacterial peptide, which may be used
CC   to produce the composition of the invention
XX
SQ   Sequence 21 AA;

  Query Match             94.2%;  Score 81;  DB 4;  Length 21;
  Best Local Similarity   87.5%;  Pred. No. 0.00052;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 DKGXXLPRPTPPRPIY 16
              |||  ||||||||||||
Db          3 DKGSYLPRPTPPRPIY 18


RESULT 19
AAY72439
ID   AAY72439 standard; peptide; 21 AA.
XX
AC   AAY72439;
XX
DT   06-AUG-2003  (revised)
DT   24-APR-2001  (first entry)
XX
DE   Pyrrhocoricin-modified Peptide 4.
XX
KW   Pyrrhocoricin-derived peptide; antibacterial; fungicidal; therapy;
KW   fungal infection; bacterial infection; candidiasis; drug development.
XX
```

OS Pyrrhocoris apterus.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Modified-site 1
FT /note= "N-terminal acetyl"
XX
PN WO200078956-A1.
XX
PD 28-DEC-2000.
XX
PF 21-JUN-2000; 2000WO-US016989.
XX
PR 23-JUN-1999; 99US-0140606P.
PR 15-SEP-1999; 99US-0154135P.
XX
PA (WIST-) WISTAR INST ANATOMY & BIOLOGY.
XX
PI Otvos L;
XX
DR WPI; 2001-112323/12.
XX
PT Polypeptides derived from the peptide pyrrhocoricin, useful for treating
PT fungal infections and Gram negative/positive bacterial infections.
XX
PS Claim 23; Page 45; 75pp; English.
XX
CC The present peptide sequence is active Pyrrhocoricin-modified Peptide 4.
CC Pyrrhocoricin is a glycopeptide characterised by the presence of a
CC disaccharide in the mid-chain position. The invention relates to
CC pyrrhocoricin-derived peptides which have anti-bacterial or anti-fungal
CC activity. These peptides have metabolic stability in mammalian serum. The
CC pyrrhocoricin-derived peptides are used in the treatment of bacterial
CC infections caused by Gram positive or Gram negative bacterium and fungal
CC infections of skin, nails, mucus membranes and intestines e.g.,
CC candidiasis. These peptides are also useful in anti-bacterial or anti-
CC fungal pharmaceutical compositions, drug development and identification
CC of other antibiotic or anti-fungal compounds. (Updated on 06-AUG-2003 to
CC correct OS field.)
XX
SQ Sequence 21 AA;

Query Match 94.2%; Score 81; DB 4; Length 21;
Best Local Similarity 87.5%; Pred. No. 0.00052;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 DKGXXLPRPTPPRPIY 16
||| |||||||||||
Db 3 DKGSYLPRPTPPRPIY 18

RESULT 20
AAY72444
ID AAY72444 standard; peptide; 21 AA.
XX
AC AAY72444;
XX
DT 06-AUG-2003 (revised)
DT 24-APR-2001 (first entry)
XX
DE Pyrrhocoricin-modified Peptide 9.
XX
KW Pyrrhocoricin-derived peptide; antibacterial; fungicidal; therapy;
KW fungal infection; bacterial infection; candidiasis; drug development.
XX
OS Pyrrhocoris apterus.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Modified-site 1
FT /note= "N-terminal acetyl"
FT Modified-site 21
FT /note= "Beta-acetyl-2,3-diamino propionic acid"
XX
PN WO200078956-A1.
XX
PD 28-DEC-2000.
XX
PF 21-JUN-2000; 2000WO-US016989.
XX
PR 23-JUN-1999; 99US-0140606P.
PR 15-SEP-1999; 99US-0154135P.
XX
PA (WIST-) WISTAR INST ANATOMY & BIOLOGY.
XX
PI Otvos L;
XX
DR WPI; 2001-112323/12.
XX
PT Polypeptides derived from the peptide pyrrhocoricin, useful for treating
PT fungal infections and Gram negative/positive bacterial infections.
XX
PS Claim 27; Page 46; 75pp; English.
XX
CC The present peptide sequence is active Pyrrhocoricin-modified Peptide 9.
CC Pyrrhocoricin is a glycopeptide characterised by the presence of a
CC disaccharide in the mid-chain position. The invention relates to
CC pyrrhocoricin-derived peptides which have anti-bacterial or anti-fungal
CC activity. These peptides have metabolic stability in mammalian serum. The
CC pyrrhocoricin-derived peptides are used in the treatment of bacterial
CC infections caused by Gram positive or Gram negative bacterium and fungal
CC infections of skin, nails, mucus membranes and intestines e.g.,
CC candidiasis. These peptides are also useful in anti-bacterial or anti-
CC fungal pharmaceutical compositions, drug development and identification
CC of other antibiotic or anti-fungal compounds. (Updated on 06-AUG-2003 to
CC correct OS field.)
XX
SQ Sequence 21 AA;

Query Match 94.2%; Score 81; DB 4; Length 21;
Best Local Similarity 87.5%; Pred. No. 0.00052;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 DKGXXLPRPTPPRPIY 16
||| |||||||||||
Db 3 DKGSYLPRPTPPRPIY 18

RESULT 21
AAY72454
ID AAY72454 standard; peptide; 21 AA.
XX
AC AAY72454;
XX
DT 06-AUG-2003 (revised)
DT 24-APR-2001 (first entry)
XX
DE Pyrrhocoricin-modified Peptide 22.
XX
KW Pyrrhocoricin-derived peptide; antibacterial; fungicidal; therapy;
KW fungal infection; bacterial infection; candidiasis; drug development.
XX
OS Pyrrhocoris apterus.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Modified-site 1
FT /note= "N-terminal acetyl"
FT Modified-site 21
FT /note= "Beta-acetyl-2,3-diamino propionic acid"
XX
PN WO200078956-A1.
XX
PD 28-DEC-2000.
XX
PF 21-JUN-2000; 2000WO-US016989.

```
XX
PR   23-JUN-1999;   99US-0140606P.
PR   15-SEP-1999;   99US-0154135P.
XX
PA   (WIST-) WISTAR INST ANATOMY & BIOLOGY.
XX
PI   Otvos L;
XX
DR   WPI; 2001-112323/12.
XX
PT   Polypeptides derived from the peptide pyrrhocoricin, useful for treating
PT   fungal infections and Gram negative/positive bacterial infections.
XX
PS   Claim 35; Page 47; 75pp; English.
XX
CC   The present peptide sequence is active Pyrrhocoricin-modified Peptide 22.
CC   Pyrrhocoricin is a glycopeptide characterised by the presence of a
CC   disaccharide in the mid-chain position. The invention relates to
CC   pyrrhocoricin-derived peptides which have anti-bacterial or anti-fungal
CC   activity. These peptides have metabolic stability in mammalian serum. The
CC   pyrrhocoricin-derived peptides are used in the treatment of bacterial
CC   infections caused by Gram positive or Gram negative bacterium and fungal
CC   infections of skin, nails, mucus membranes and intestines e.g.,
CC   candidiasis. These peptides are also useful in anti-bacterial or anti-
CC   fungal pharmaceutical compositions, drug development and identification
CC   of other antibiotic or anti-fungal compounds. (Updated on 06-AUG-2003 to
CC   correct OS field.)
XX
SQ   Sequence 21 AA;

  Query Match             94.2%;  Score 81;  DB 4;  Length 21;
  Best Local Similarity   87.5%;  Pred. No. 0.00052;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 DKGXXLPRPTPPRPIY 16
              |||  ||||||||||
Db          3 DKGSYLPRPTPPRPIY 18


RESULT 22
AAY72448
ID   AAY72448 standard; peptide; 21 AA.
XX
AC   AAY72448;
XX
DT   06-AUG-2003  (revised)
DT   24-APR-2001  (first entry)
XX
DE   Pyrrhocoricin-modified Peptide 16.
XX
KW   Pyrrhocoricin-derived peptide; antibacterial; fungicidal; therapy;
KW   fungal infection; bacterial infection; candidiasis; drug development;
KW   cyclic.
XX
OS   Pyrrhocoris apterus.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   1
FT                   /note= "Forms a cyclic linkage with Asp at the C-terminal
FT                   end"
FT   Modified-site   21
FT                   /note= "Forms a cyclic linkage with Lys at the N-terminal
FT                   end"
XX
PN   WO200078956-A1.
XX
PD   28-DEC-2000.
XX
PF   21-JUN-2000; 2000WO-US016989.
XX
PR   23-JUN-1999;   99US-0140606P.
PR   15-SEP-1999;   99US-0154135P.
XX
PA   (WIST-) WISTAR INST ANATOMY & BIOLOGY.
XX
PI   Otvos L;
XX
DR   WPI; 2001-112323/12.
XX
PT   Polypeptides derived from the peptide pyrrhocoricin, useful for treating
PT   fungal infections and Gram negative/positive bacterial infections.
XX
PS   Example 1; Page 26; 75pp; English.
XX
CC   The present peptide sequence is inactive Pyrrhocoricin-modified Peptide
CC   16. Pyrrhocoricin is a glycopeptide characterised by the presence of a
CC   disaccharide in the mid-chain position. The invention relates to
CC   pyrrhocoricin-derived peptides which have anti-bacterial or anti-fungal
CC   activity. These peptides have metabolic stability in mammalian serum. The
CC   pyrrhocoricin-derived peptides are used in the treatment of bacterial
CC   infections caused by Gram positive or Gram negative bacterium and fungal
CC   infections of skin, nails, mucus membranes and intestines e.g.,
CC   candidiasis. These peptides are also useful in anti-bacterial or anti-
CC   fungal pharmaceutical compositions, drug development and identification
CC   of other antibiotic or anti-fungal compounds. (Updated on 06-AUG-2003 to
CC   correct OS field.)
XX
SQ   Sequence 21 AA;

  Query Match             94.2%;  Score 81;  DB 4;  Length 21;
  Best Local Similarity   87.5%;  Pred. No. 0.00052;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 DKGXXLPRPTPPRPIY 16
              |||  ||||||||||
Db          3 DKGSYLPRPTPPRPIY 18


RESULT 23
AAY72440
ID   AAY72440 standard; peptide; 21 AA.
XX
AC   AAY72440;
XX
DT   06-AUG-2003  (revised)
DT   24-APR-2001  (first entry)
XX
DE   Pyrrhocoricin-modified Peptide 5.
XX
KW   Pyrrhocoricin-derived peptide; antibacterial; fungicidal; therapy;
KW   fungal infection; bacterial infection; candidiasis; drug development.
XX
OS   Pyrrhocoris apterus.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   1
FT                   /note= "N-terminal acetyl"
XX
PN   WO200078956-A1.
XX
PD   28-DEC-2000.
XX
PF   21-JUN-2000; 2000WO-US016989.
XX
PR   23-JUN-1999;   99US-0140606P.
PR   15-SEP-1999;   99US-0154135P.
XX
PA   (WIST-) WISTAR INST ANATOMY & BIOLOGY.
XX
PI   Otvos L;
XX
DR   WPI; 2001-112323/12.
XX
PT   Polypeptides derived from the peptide pyrrhocoricin, useful for treating
```

PT fungal infections and Gram negative/positive bacterial infections.
XX
PS Claim 24; Page 45; 75pp; English.
XX
CC The present peptide sequence is active Pyrrhocoricin-modified Peptide 5.
CC Pyrrhocoricin is a glycopeptide characterised by the presence of a
CC disaccharide in the mid-chain position. The invention relates to
CC pyrrhocoricin-derived peptides which have anti-bacterial or anti-fungal
CC activity. These peptides have metabolic stability in mammalian serum. The
CC pyrrhocoricin-derived peptides are used in the treatment of bacterial
CC infections caused by Gram positive or Gram negative bacterium and fungal
CC infections of skin, nails, mucus membranes and intestines e.g.,
CC candidiasis. These peptides are also useful in anti-bacterial or anti-
CC fungal pharmaceutical compositions, drug development and identification
CC of other antibiotic or anti-fungal compounds. (Updated on 06-AUG-2003 to
CC correct OS field.)
XX
SQ Sequence 21 AA;

Query Match 94.2%; Score 81; DB 4; Length 21;
Best Local Similarity 87.5%; Pred. No. 0.00052;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 DKGXXLPRPTPPRPIY 16
||| |||||||||||
Db 3 DKGSYLPRPTPPRPIY 18

RESULT 24
AAY72451
ID AAY72451 standard; peptide; 21 AA.
XX
AC AAY72451;
XX
DT 06-AUG-2003 (revised)
DT 24-APR-2001 (first entry)
XX
DE Pyrrhocoricin-modified Peptide 19.
XX
KW Pyrrhocoricin-derived peptide; antibacterial; fungicidal; therapy;
KW fungal infection; bacterial infection; candidiasis; drug development.
XX
OS Pyrrhocoris apterus.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Modified-site 1
FT /note= "N-terminal 5(6)-carboxyfluorescein"
XX
PN WO200078956-A1.
XX
PD 28-DEC-2000.
XX
PF 21-JUN-2000; 2000WO-US016989.
XX
PR 23-JUN-1999; 99US-0140606P.
PR 15-SEP-1999; 99US-0154135P.
XX
PA (WIST-) WISTAR INST ANATOMY & BIOLOGY.
XX
PI Otvos L;
XX
DR WPI; 2001-112323/12.
XX
PT Polypeptides derived from the peptide pyrrhocoricin, useful for treating
PT fungal infections and Gram negative/positive bacterial infections.
XX
PS Claim 32; Page 46; 75pp; English.
XX
CC The present peptide sequence is active Pyrrhocoricin-modified Peptide 19.
CC Pyrrhocoricin is a glycopeptide characterised by the presence of a
CC disaccharide in the mid-chain position. The invention relates to
CC pyrrhocoricin-derived peptides which have anti-bacterial or anti-fungal
CC activity. These peptides have metabolic stability in mammalian serum. The
CC pyrrhocoricin-derived peptides are used in the treatment of bacterial
CC infections caused by Gram positive or Gram negative bacterium and fungal
CC infections of skin, nails, mucus membranes and intestines e.g.,
CC candidiasis. These peptides are also useful in anti-bacterial or anti-
CC fungal pharmaceutical compositions, drug development and identification
CC of other antibiotic or anti-fungal compounds. (Updated on 06-AUG-2003 to
CC correct OS field.)
XX
SQ Sequence 21 AA;

Query Match 94.2%; Score 81; DB 4; Length 21;
Best Local Similarity 87.5%; Pred. No. 0.00052;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 DKGXXLPRPTPPRPIY 16
||| |||||||||||
Db 3 DKGSYLPRPTPPRPIY 18

RESULT 25
AAY72452
ID AAY72452 standard; peptide; 21 AA.
XX
AC AAY72452;
XX
DT 06-AUG-2003 (revised)
DT 24-APR-2001 (first entry)
XX
DE Pyrrhocoricin-modified Peptide 20.
XX
KW Pyrrhocoricin-derived peptide; antibacterial; fungicidal; therapy;
KW fungal infection; bacterial infection; candidiasis; drug development.
XX
OS Pyrrhocoris apterus.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Modified-site 1
FT /note= "N-terminal acetyl"
FT Misc-difference 21
FT /note= "Wild type Asn substituted with Asp"
XX
PN WO200078956-A1.
XX
PD 28-DEC-2000.
XX
PF 21-JUN-2000; 2000WO-US016989.
XX
PR 23-JUN-1999; 99US-0140606P.
PR 15-SEP-1999; 99US-0154135P.
XX
PA (WIST-) WISTAR INST ANATOMY & BIOLOGY.
XX
PI Otvos L;
XX
DR WPI; 2001-112323/12.
XX
PT Polypeptides derived from the peptide pyrrhocoricin, useful for treating
PT fungal infections and Gram negative/positive bacterial infections.
XX
PS Claim 33; Page 46; 75pp; English.
XX
CC The present peptide sequence is weakly active Pyrrhocoricin-modified
CC Peptide 20. Pyrrhocoricin is a glycopeptide characterised by the presence
CC of a disaccharide in the mid-chain position. The invention relates to
CC pyrrhocoricin-derived peptides which have anti-bacterial or anti-fungal
CC activity. These peptides have metabolic stability in mammalian serum. The
CC pyrrhocoricin-derived peptides are used in the treatment of bacterial
CC infections caused by Gram positive or Gram negative bacterium and fungal
CC infections of skin, nails, mucus membranes and intestines e.g.,
CC candidiasis. These peptides are also useful in anti-bacterial or anti-
CC fungal pharmaceutical compositions, drug development and identification

```
CC   of other antibiotic or anti-fungal compounds. (Updated on 06-AUG-2003 to
CC   correct OS field.)
XX
SQ   Sequence 21 AA;

  Query Match             94.2%;  Score 81;  DB 4;  Length 21;
  Best Local Similarity   87.5%;  Pred. No. 0.00052;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 DKGXXLPRPTPPRPIY 16
              |||  ||||||||||||
Db          3 DKGSYLPRPTPPRPIY 18


RESULT 26
AAY72450
ID   AAY72450 standard; peptide; 21 AA.
XX
AC   AAY72450;
XX
DT   06-AUG-2003  (revised)
DT   24-APR-2001  (first entry)
XX
DE   Pyrrhocoricin-modified peptide 18.
XX
KW   Pyrrhocoricin-derived peptide; antibacterial; fungicidal; therapy;
KW   fungal infection; bacterial infection; candidiasis; drug development.
XX
OS   Pyrrhocoris apterus.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   1
FT                   /note= "N-terminal biotin"
XX
PN   WO200078956-A1.
XX
PD   28-DEC-2000.
XX
PF   21-JUN-2000; 2000WO-US016989.
XX
PR   23-JUN-1999;   99US-0140606P.
PR   15-SEP-1999;   99US-0154135P.
XX
PA   (WIST-) WISTAR INST ANATOMY & BIOLOGY.
XX
PI   Otvos L;
XX
DR   WPI; 2001-112323/12.
XX
PT   Polypeptides derived from the peptide pyrrhocoricin, useful for treating
PT   fungal infections and Gram negative/positive bacterial infections.
XX
PS   Claim 31; Page 46; 75pp; English.
XX
CC   The present peptide sequence is active Pyrrhocoricin-modified Peptide 18.
CC   Pyrrhocoricin is a glycopeptide characterised by the presence of a
CC   disaccharide in the mid-chain position. The invention relates to
CC   pyrrhocoricin-derived peptides which have anti-bacterial or anti-fungal
CC   activity. These peptides have metabolic stability in mammalian serum. The
CC   pyrrhocoricin-derived peptides are used in the treatment of bacterial
CC   infections caused by Gram positive or Gram negative bacterium and fungal
CC   infections of skin, nails, mucus membranes and intestines e.g.,
CC   candidiasis. These peptides are also useful in anti-bacterial or anti-
CC   fungal pharmaceutical compositions, drug development and identification
CC   of other antibiotic or anti-fungal compounds. (Updated on 06-AUG-2003 to
CC   correct OS field.)
XX
SQ   Sequence 21 AA;

  Query Match             94.2%;  Score 81;  DB 4;  Length 21;
  Best Local Similarity   87.5%;  Pred. No. 0.00052;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 DKGXXLPRPTPPRPIY 16
              |||  ||||||||||||
Db          3 DKGSYLPRPTPPRPIY 18


RESULT 27
AAY72445
ID   AAY72445 standard; peptide; 21 AA.
XX
AC   AAY72445;
XX
DT   06-AUG-2003  (revised)
DT   24-APR-2001  (first entry)
XX
DE   Pyrrhocoricin-modified Peptide 10.
XX
KW   Pyrrhocoricin-derived peptide; antibacterial; fungicidal; therapy;
KW   fungal infection; bacterial infection; candidiasis; drug development.
XX
OS   Pyrrhocoris apterus.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   1
FT                   /note= "N-terminal acetyl"
FT   Modified-site   21
FT                   /note= "Modified with 2-acetamido-2-deoxyglucose
FT                   (GlcNAc)"
XX
PN   WO200078956-A1.
XX
PD   28-DEC-2000.
XX
PF   21-JUN-2000; 2000WO-US016989.
XX
PR   23-JUN-1999;   99US-0140606P.
PR   15-SEP-1999;   99US-0154135P.
XX
PA   (WIST-) WISTAR INST ANATOMY & BIOLOGY.
XX
PI   Otvos L;
XX
DR   WPI; 2001-112323/12.
XX
PT   Polypeptides derived from the peptide pyrrhocoricin, useful for treating
PT   fungal infections and Gram negative/positive bacterial infections.
XX
PS   Claim 28; Page 46; 75pp; English.
XX
CC   The present peptide sequence is active Pyrrhocoricin-modified Peptide 10.
CC   Pyrrhocoricin is a glycopeptide characterised by the presence of a
CC   disaccharide in the mid-chain position. The invention relates to
CC   pyrrhocoricin-derived peptides which have anti-bacterial or anti-fungal
CC   activity. These peptides have metabolic stability in mammalian serum. The
CC   pyrrhocoricin-derived peptides are used in the treatment of bacterial
CC   infections caused by Gram positive or Gram negative bacterium and fungal
CC   infections of skin, nails, mucus membranes and intestines e.g.,
CC   candidiasis. These peptides are also useful in anti-bacterial or anti-
CC   fungal pharmaceutical compositions, drug development and identification
CC   of other antibiotic or anti-fungal compounds. (Updated on 06-AUG-2003 to
CC   correct OS field.)
XX
SQ   Sequence 21 AA;

  Query Match             94.2%;  Score 81;  DB 4;  Length 21;
  Best Local Similarity   87.5%;  Pred. No. 0.00052;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 DKGXXLPRPTPPRPIY 16
              |||  ||||||||||||
Db          3 DKGSYLPRPTPPRPIY 18
```

```
RESULT 28
AAY72446
ID   AAY72446 standard; peptide; 21 AA.
XX
AC   AAY72446;
XX
DT   06-AUG-2003  (revised)
DT   24-APR-2001  (first entry)
XX
DE   Pyrrhocoricin-modified Peptide 11.
XX
KW   Pyrrhocoricin-derived peptide; antibacterial; fungicidal; therapy;
KW   fungal infection; bacterial infection; candidiasis; drug development.
XX
OS   Pyrrhocoris apterus.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   1
FT                   /note= "N-terminal acetyl"
FT   Modified-site   21
FT                   /note= "Modified with triacetyl-2-acetamido-2-
FT                   deoxyglucose (Ac3-GlcNAc)"
XX
PN   WO200078956-A1.
XX
PD   28-DEC-2000.
XX
PF   21-JUN-2000; 2000WO-US016989.
XX
PR   23-JUN-1999;   99US-0140606P.
PR   15-SEP-1999;   99US-0154135P.
XX
PA   (WIST-) WISTAR INST ANATOMY & BIOLOGY.
XX
PI   Otvos L;
XX
DR   WPI; 2001-112323/12.
XX
PT   Polypeptides derived from the peptide pyrrhocoricin, useful for treating
PT   fungal infections and Gram negative/positive bacterial infections.
XX
PS   Claim 29; Page 46; 75pp; English.
XX
CC   The present peptide sequence is active Pyrrhocoricin-modified Peptide 11.
CC   Pyrrhocoricin is a glycopeptide characterised by the presence of a
CC   disaccharide in the mid-chain position. The invention relates to
CC   pyrrhocoricin-derived peptides which have anti-bacterial or anti-fungal
CC   activity. These peptides have metabolic stability in mammalian serum. The
CC   pyrrhocoricin-derived peptides are used in the treatment of bacterial
CC   infections caused by Gram positive or Gram negative bacterium and fungal
CC   infections of skin, nails, mucus membranes and intestines e.g.,
CC   candidiasis. These peptides are also useful in anti-bacterial or anti-
CC   fungal pharmaceutical compositions, drug development and identification
CC   of other antibiotic or anti-fungal compounds. (Updated on 06-AUG-2003 to
CC   correct OS field.)
XX
SQ   Sequence 21 AA;

  Query Match             94.2%;  Score 81;  DB 4;  Length 21;
  Best Local Similarity   87.5%;  Pred. No. 0.00052;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 DKGXXLPRPTPPRPIY 16
              |||  ||||||||||||
Db          3 DKGSYLPRPTPPRPIY 18


RESULT 29
AAY72461
ID   AAY72461 standard; peptide; 23 AA.
XX
AC   AAY72461;
XX
DT   06-AUG-2003  (revised)
DT   24-APR-2001  (first entry)
XX
DE   Pyrrhocoricin-modified peptide.
XX
KW   Pyrrhocoricin-derived peptide; antibacterial; fungicidal; therapy;
KW   fungal infection; bacterial infection; candidiasis; drug development.
XX
OS   Pyrrhocoris apterus.
OS   Synthetic.
XX
PN   WO200078956-A1.
XX
PD   28-DEC-2000.
XX
PF   21-JUN-2000; 2000WO-US016989.
XX
PR   23-JUN-1999;   99US-0140606P.
PR   15-SEP-1999;   99US-0154135P.
XX
PA   (WIST-) WISTAR INST ANATOMY & BIOLOGY.
XX
PI   Otvos L;
XX
DR   WPI; 2001-112323/12.
XX
PT   Polypeptides derived from the peptide pyrrhocoricin, useful for treating
PT   fungal infections and Gram negative/positive bacterial infections.
XX
PS   Claim 21; Page 45; 75pp; English.
XX
CC   The present peptide sequence is inactive Pyrrhocoricin-modified peptide.
CC   Pyrrhocoricin is a glycopeptide characterised by the presence of a
CC   disaccharide in the mid-chain position. The invention relates to
CC   pyrrhocoricin-derived peptides which have anti-bacterial or anti-fungal
CC   activity. These peptides have metabolic stability in mammalian serum. The
CC   pyrrhocoricin-derived peptides are used in the treatment of bacterial
CC   infections caused by Gram positive or Gram negative bacterium and fungal
CC   infections of skin, nails, mucus membranes and intestines e.g.,
CC   candidiasis. These peptides are also useful in anti-bacterial or anti-
CC   fungal pharmaceutical compositions, drug development and identification
CC   of other antibiotic or anti-fungal compounds. (Updated on 06-AUG-2003 to
CC   correct OS field.)
XX
SQ   Sequence 23 AA;

  Query Match             94.2%;  Score 81;  DB 4;  Length 23;
  Best Local Similarity   87.5%;  Pred. No. 0.00056;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 DKGXXLPRPTPPRPIY 16
              |||  ||||||||||||
Db          5 DKGSYLPRPTPPRPIY 20


RESULT 30
AAY72438
ID   AAY72438 standard; peptide; 24 AA.
XX
AC   AAY72438;
XX
DT   06-AUG-2003  (revised)
DT   24-APR-2001  (first entry)
XX
DE   Pyrrhocoricin-modified Peptide 3.
XX
KW   Pyrrhocoricin-derived peptide; antibacterial; fungicidal; therapy;
KW   fungal infection; bacterial infection; candidiasis; drug development.
XX
OS   Pyrrhocoris apterus.
OS   Synthetic.
```

```
XX
FH   Key             Location/Qualifiers
FT   Modified-site   1
FT                   /note= "N-terminal acetyl"
XX
PN   WO200078956-A1.
XX
PD   28-DEC-2000.
XX
PF   21-JUN-2000; 2000WO-US016989.
XX
PR   23-JUN-1999;   99US-0140606P.
PR   15-SEP-1999;   99US-0154135P.
XX
PA   (WIST-) WISTAR INST ANATOMY & BIOLOGY.
XX
PI   Otvos L;
XX
DR   WPI; 2001-112323/12.
XX
PT   Polypeptides derived from the peptide pyrrhocoricin, useful for treating
PT   fungal infections and Gram negative/positive bacterial infections.
XX
PS   Claim 22; Page 45; 75pp; English.
XX
CC   The present peptide sequence is active Pyrrhocoricin-modified Peptide 3.
CC   Pyrrhocoricin is a glycopeptide characterised by the presence of a
CC   disaccharide in the mid-chain position. The invention relates to
CC   pyrrhocoricin-derived peptides which have anti-bacterial or anti-fungal
CC   activity. These peptides have metabolic stability in mammalian serum. The
CC   pyrrhocoricin-derived peptides are used in the treatment of bacterial
CC   infections caused by Gram positive or Gram negative bacterium and fungal
CC   infections of skin, nails, mucus membranes and intestines e.g.,
CC   candidiasis. These peptides are also useful in anti-bacterial or anti-
CC   fungal pharmaceutical compositions, drug development and identification
CC   of other antibiotic or anti-fungal compounds. (Updated on 06-AUG-2003 to
CC   correct OS field.)
XX
SQ   Sequence 24 AA;

  Query Match             94.2%;  Score 81;  DB 4;  Length 24;
  Best Local Similarity   87.5%;  Pred. No. 0.00059;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 DKGXXLPRPTPPRPIY 16
              |||  ||||||||||
Db          6 DKGSYLPRPTPPRPIY 21


RESULT 31
AAY72449
ID   AAY72449 standard; peptide; 29 AA.
XX
AC   AAY72449;
XX
DT   06-AUG-2003  (revised)
DT   24-APR-2001  (first entry)
XX
DE   Pyrrhocoricin-modified Peptide 17.
XX
KW   Pyrrhocoricin-derived peptide; antibacterial; fungicidal; therapy;
KW   fungal infection; bacterial infection; candidiasis; drug development;
KW   cyclic.
XX
OS   Pyrrhocoris apterus.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   1
FT                   /note= "Forms a cyclic linkage with Asn at the C-terminal
FT                   end"
FT   Modified-site   29
FT                   /note= "Forms a cyclic linkage with Arg at the N-terminal
FT                   end"
XX
PN   WO200078956-A1.
XX
PD   28-DEC-2000.
XX
PF   21-JUN-2000; 2000WO-US016989.
XX
PR   23-JUN-1999;   99US-0140606P.
PR   15-SEP-1999;   99US-0154135P.
XX
PA   (WIST-) WISTAR INST ANATOMY & BIOLOGY.
XX
PI   Otvos L;
XX
DR   WPI; 2001-112323/12.
XX
PT   Polypeptides derived from the peptide pyrrhocoricin, useful for treating
PT   fungal infections and Gram negative/positive bacterial infections.
XX
PS   Claim 37; Page 47; 75pp; English.
XX
CC   The present peptide sequence is active Pyrrhocoricin-modified Peptide 17.
CC   This cyclic non-glycosylated peptide is the most active peptide.
CC   Pyrrhocoricin is a glycopeptide characterised by the presence of a
CC   disaccharide in the mid-chain position. The invention relates to
CC   pyrrhocoricin-derived peptides which have anti-bacterial or anti-fungal
CC   activity. These peptides have metabolic stability in mammalian serum. The
CC   pyrrhocoricin-derived peptides are used in the treatment of bacterial
CC   infections caused by Gram positive or Gram negative bacterium and fungal
CC   infections of skin, nails, mucus membranes and intestines e.g.,
CC   candidiasis. These peptides are also useful in anti-bacterial or anti-
CC   fungal pharmaceutical compositions, drug development and identification
CC   of other antibiotic or anti-fungal compounds. (Updated on 06-AUG-2003 to
CC   correct OS field.)
XX
SQ   Sequence 29 AA;

  Query Match             94.2%;  Score 81;  DB 4;  Length 29;
  Best Local Similarity   87.5%;  Pred. No. 0.00069;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 DKGXXLPRPTPPRPIY 16
              |||  ||||||||||
Db         11 DKGSYLPRPTPPRPIY 26


RESULT 32
AAG62740
ID   AAG62740 standard; peptide; 18 AA.
XX
AC   AAG62740;
XX
DT   17-SEP-2001  (first entry)
XX
DE   Amino acid sequence of modified antibacterial peptide pyrrhocoricin.
XX
KW   Multi-helical lid; heat shock protein; hsp; protein folding;
KW   pathogenic infection; bacterial infection; antibacterial.
XX
OS   Unidentified.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   1
FT                   /note= "a moiety having a net positive charge is
FT                   attached"
XX
PN   WO200153509-A2.
XX
PD   26-JUL-2001.
XX
PF   19-JAN-2001; 2001WO-US001812.
XX
```

```
PR   21-JAN-2000; 2000US-0177565P.
PR   03-OCT-2000; 2000US-0237599P.
XX
PA   (WIST-) WISTAR INST ANATOMY & BIOLOGY.
PA   (UYCR-) UNIV CREIGHTON.
XX
PI   Otvos L,  Blaszczyk-Thurin M,  Rogers M,  Lovas S;
XX
DR   WPI; 2001-451911/48.
XX
PT   Composition, used to treat a pathogenic infection and eliminate a plant,
PT   insect, or animal pest, comprises a molecule that binds to a heat shock
PT   protein.
XX
PS   Disclosure; Page 111; 124pp; English.
XX
CC   The specification describes a composition that comprises a synthetic non-
CC   naturally occurring molecule that binds to a selected multi-helical lid
CC   of a heat shock protein (hsp) of a selected organism, where the molecule
CC   inhibits protein folding activity of the hsp, and a carrier, where
CC   exposure of the organism to the composition retards the growth and
CC   reproduction of the organism. The composition is used to treat a mammal
CC   suffering from a pathogenic infection, in the manufacture of a medicament
CC   for treating a mammal for a pathogenic infection, and to eliminate a
CC   plant, insect, or animal pest. It is used in the manufacture of a
CC   medicament for treating mammalian bacterial infection. The present
CC   sequence represents a modified antibacterial peptide, which may be used
CC   to produce the composition of the invention
XX
SQ   Sequence 18 AA;

  Query Match             93.0%;  Score 80;  DB 4;  Length 18;
  Best Local Similarity   100.0%;  Pred. No. 0.00061;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 DKGXXLPRPTPPRPIY 16
              ||||||||||||||||
Db          1 DKGXXLPRPTPPRPIY 16


RESULT 33
AAY72424
ID   AAY72424 standard; peptide; 18 AA.
XX
AC   AAY72424;
XX
DT   06-AUG-2003  (revised)
DT   24-APR-2001  (first entry)
XX
DE   Pyrrhocoricin based generic peptide #1.
XX
KW   Pyrrhocoricin-derived peptide; antibacterial; fungicidal; therapy;
KW   fungal infection; bacterial infection; candidiasis; drug development.
XX
OS   Pyrrhocoris apterus.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   1
FT                   /note= "Optionally attached to additional amino acids or
FT                   modified with a straight chain, branched, cyclic or
FT                   heterocyclic alkyl group (preferably 1-aminocyclo-hexane
FT                   carboxylic acid), heterocyclic alkanoyl group or a
FT                   positively charged reporter group (preferably biotin,
FT                   5(6) carboxyfluorescein)"
FT   Misc-difference 4
FT                   /note= "Ser or any amino acid"
FT   Misc-difference 5
FT                   /note= "Tyr or any amino acid"
FT   Misc-difference 17
FT                   /note= "Asn or any amino acid"
FT   Modified-site   18
FT                   /note= "Optionally attached to additional amino acids or
FT                   modified with an amide, an imide or a sugar moiety"
FT   Misc-difference 18
FT                   /note= "Arg or any amino acid"
XX
PN   WO200078956-A1.
XX
PD   28-DEC-2000.
XX
PF   21-JUN-2000; 2000WO-US016989.
XX
PR   23-JUN-1999;   99US-0140606P.
PR   15-SEP-1999;   99US-0154135P.
XX
PA   (WIST-) WISTAR INST ANATOMY & BIOLOGY.
XX
PI   Otvos L;
XX
DR   WPI; 2001-112323/12.
XX
PT   Polypeptides derived from the peptide pyrrhocoricin, useful for treating
PT   fungal infections and Gram negative/positive bacterial infections.
XX
PS   Claim 1; Page 42; 75pp; English.
XX
CC   The present sequence is a pyrrhocoricin based generic peptide which has
CC   anti-bacterial or anti-fungal activity. Pyrrhocoricin is a glycopeptide
CC   characterised by the presence of a disaccharide in the mid-chain
CC   position. The invention relates to pyrrhocoricin-derived peptides. These
CC   peptides have metabolic stability in mammalian serum. The pyrrhocoricin-
CC   derived peptides are used in the treatment of bacterial infections caused
CC   by Gram positive or Gram negative bacterium and fungal infections of
CC   skin, nails, mucus membranes and intestines e.g., candidiasis. These
CC   peptides are also useful in anti-bacterial or anti-fungal pharmaceutical
CC   compositions, drug development and identification of other antibiotic or
CC   anti-fungal compounds. (Updated on 06-AUG-2003 to correct OS field.)
XX
SQ   Sequence 18 AA;

  Query Match             93.0%;  Score 80;  DB 4;  Length 18;
  Best Local Similarity   100.0%;  Pred. No. 0.00061;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 DKGXXLPRPTPPRPIY 16
              ||||||||||||||||
Db          1 DKGXXLPRPTPPRPIY 16


RESULT 34
AAG62767
ID   AAG62767 standard; peptide; 18 AA.
XX
AC   AAG62767;
XX
DT   17-SEP-2001  (first entry)
XX
DE   Amino acid sequence of modified antibacterial peptide pyrrhocoricin.
XX
KW   Multi-helical lid; heat shock protein; hsp; protein folding;
KW   pathogenic infection; bacterial infection; antibacterial.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   1
FT                   /note= "1-aminocyclo-hexane carboxylic"
XX
PN   WO200153509-A2.
XX
PD   26-JUL-2001.
XX
PF   19-JAN-2001; 2001WO-US001812.
XX
PR   21-JAN-2000; 2000US-0177565P.
```

PR 03-OCT-2000; 2000US-0237599P.
XX
PA (WIST-) WISTAR INST ANATOMY & BIOLOGY.
PA (UYCR-) UNIV CREIGHTON.
XX
PI Otvos L, Blaszczyk-Thurin M, Rogers M, Lovas S;
XX
DR WPI; 2001-451911/48.
XX
PT Composition, used to treat a pathogenic infection and eliminate a plant,
PT insect, or animal pest, comprises a molecule that binds to a heat shock
PT protein.
XX
PS Example 4; Page 62; 124pp; English.
XX
CC The specification describes a composition that comprises a synthetic non-
CC naturally occurring molecule that binds to a selected multi-helical lid
CC of a heat shock protein (hsp) of a selected organism, where the molecule
CC inhibits protein folding activity of the hsp, and a carrier, where
CC exposure of the organism to the composition retards the growth and
CC reproduction of the organism. The composition is used to treat a mammal
CC suffering from a pathogenic infection, in the manufacture of a medicament
CC for treating a mammal for a pathogenic infection, and to eliminate a
CC plant, insect, or animal pest. It is used in the manufacture of a
CC medicament for treating mammalian bacterial infection. The present
CC sequence represents a modified antibacterial peptide, which may be used
CC to produce the composition of the invention
XX
SQ Sequence 18 AA;

Query Match 86.0%; Score 74; DB 4; Length 18;
Best Local Similarity 81.2%; Pred. No. 0.0037;
Matches 13; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 DKGXXLPRPTPPRPIY 16
|| | ||||||||||
Db 1 DLGSYLPRPTPPRPIY 16

RESULT 35
ABO00915
ID ABO00915 standard; protein; 133 AA.
XX
AC ABO00915;
XX
DT 06-AUG-2003 (first entry)
XX
DE Polypeptide encoded by novel human contig #166.
XX
KW Human; angiogenesis; cytokine; cell proliferation; pluripotent;
KW cell differentiation; totipotent; stem cell; transplantation; bio-sensor;
KW neuroepithelial cell; autoimmune disease; neural cell; genetic disorder;
KW nerve; brain tissue; central nervous system disease;
KW peripheral nervous system disease; neuropathy; haematopoiesis; bone;
KW myeloid disorder; lymphoid cell disorder; platelet disorder; tendon;
KW regeneration; cartilage; tendon; ligament; nerve tissue growth;
KW tissue repair; wound healing; burn; ulcer; osteoporosis; cancer;
KW osteoarthritis; bone degenerative disorder; periodontal disease;
KW gut protection; lung fibrosis; liver fibrosis; reperfusion injury;
KW immune deficiency; infection; autoimmune disorder; allergic reaction;
KW thrombolysis; thrombosis; coagulation disorder; hereditary disorder;
KW biorhythm; circadian cycle; fertility; metabolism; catabolism; anabolism;
KW nootropic; neuroprotective; antiparkinsonian; anticonvulsant;
KW haemostatic; vulnerary; antiulcer; osteopathic; antiarthritic;
KW vasotropic; immunostimulant; antibacterial; fungicide; immunosuppressive;
KW antirheumatic; antidiabetic; antiasthmatic; cytostatic; virucide.
XX
OS Homo sapiens.
XX
PN WO2003023013-A2.
XX
PD 20-MAR-2003.
XX
PF 13-SEP-2002; 2002WO-US029001.
XX
PR 13-SEP-2001; 2001US-0322511P.
PR 12-SEP-2002; 2002US-00243552.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Yang Y, Wang Z, Weng G, Ma Y;
XX
DR WPI; 2003-313249/30.
DR N-PSDB; ACD05992.
XX
PT Novel nucleic acids and polypeptides for diagnosis, treatment of central
PT and peripheral nervous system diseases and neuropathies, such as
PT Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
PT lateral sclerosis.
XX
PS Example 3; SEQ ID NO 1039; 300pp; English.
XX
CC The present invention relates to the isolation of novel human
CC polynucleotide sequences and their encoding polypeptides. The novel
CC polypeptides exhibit activities relating to angiogenesis, cytokine, cell
CC proliferation, cell differentiation, antiinflammatory, and stem cell
CC growth factor activities. The polypeptides are involved in the
CC proliferation, differentiation and survival of pluripotent and totipotent
CC stem cells, and are useful for re-engineering damaged or diseased
CC tissues, transplantation, manufacture of bio-pharmaceuticals and
CC development of bio-sensors. The polypeptides can be used to manipulate
CC stem cells in culture to give rise to neuroepithelial cells that can be
CC used to augment or replace cells damaged by illness, autoimmune disease,
CC accidental damage or genetic disorders. The polypeptides induce the
CC proliferation of neural cells and regeneration of nerve and brain tissue
CC and are useful for the treatment of central and peripheral nervous system
CC diseases and neuropathies, such as Alzheimer's, Parkinson's disease,
CC Huntington's disease, amyotrophic lateral sclerosis (ALS). The
CC polypeptides are also involved in chemotactic or chemokinetic activity,
CC regulation of haematopoiesis and are useful for treating myeloid or
CC lymphoid cell disorders, platelet disorders such as thrombocytopaenia and
CC for regeneration of bone, cartilage, tendon, ligament and/or nerve tissue
CC growth, in tissue repair, healing of burns, incisions, ulcers, for
CC treating osteoporosis, osteoarthritis, bone degenerative disorders, and
CC periodontal disease. The polypeptides are also useful for gut protection
CC or regeneration and treatment of lung or liver fibrosis, reperfusion
CC injury in various tissues, various immune deficiencies and disorders
CC including severe combined immunodeficiency (SCID), bacterial or fungal
CC infections, autoimmune disorders (e.g. multiple sclerosis, rheumatoid
CC arthritis, diabetes mellitus, myasthenia gravis), allergic reactions and
CC conditions, such as asthma or other respiratory problems. The
CC polypeptides are involved in thrombolysis or thrombosis and are useful in
CC treatment of various coagulation disorders (including hereditary
CC disorders such as haemophilia) or to enhance coagulation and other
CC haemostatic events in treating wounds resulting from trauma, surgery or
CC other causes. The polypeptides exhibit immune stimulating or immune
CC suppressing activity, and are useful for treating autoimmune diseases or
CC cancer. They also inhibit the growth, infection or function of infectious
CC agents such as bacteria, fungi, viruses, effect biorhythms or circadian
CC cycles of rhythms, fertility of male or female subjects, metabolism,
CC catabolism, and anabolism. ABO00750-ABO00950 represent polypeptides
CC encoded by novel contigs assembled in the examples of the present
CC invention. Note: The sequence data for this patent did not form part of
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 133 AA;

Query Match 60.5%; Score 52; DB 6; Length 133;
Best Local Similarity 90.0%; Pred. No. 17;
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 6 LPRPTPPRPI 15
|||| |||||
Db 79 LPRPLPPRPI 88

Search completed: March 11, 2004, 17:18:19
Job time : 55 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: March 11, 2004, 16:55:31 ; Search time 34 Seconds
(without alignments)
111.787 Million cell updates/sec

Title: US-09-980-804-1
Perfect score: 86
Sequence: 1 DKGXXLPRPTPPRPIYXX 18

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 809742 seqs, 211153259 residues

Total number of hits satisfying chosen parameters: 809742

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	81	94.2	20	14	US-10-181-654-3	Sequence 3, Appli
2	81	94.2	21	14	US-10-181-654-12	Sequence 12, Appl
3	81	94.2	21	14	US-10-181-654-25	Sequence 25, Appl
4	80	93.0	18	14	US-10-181-654-9	Sequence 9, Appli
5	74	86.0	18	14	US-10-181-654-36	Sequence 36, Appl
6	47	54.7	176	9	US-09-953-342-25	Sequence 25, Appl
7	47	54.7	392	14	US-10-156-761-11324	Sequence 11324, A
8	47	54.7	1126	15	US-10-108-260A-3665	Sequence 3665, Ap
9	46	53.5	20	14	US-10-181-654-7	Sequence 7, Appli
10	46	53.5	487	14	US-10-224-999A-3465	Sequence 3465, Ap
11	46	53.5	3298	14	US-10-149-819-21	Sequence 21, Appl
12	46	53.5	3301	16	US-10-038-854-68	Sequence 68, Appl
13	46	53.5	3312	14	US-10-225-567A-656	Sequence 656, App
14	46	53.5	3312	16	US-10-038-854-67	Sequence 67, Appl
15	46	53.5	3313	9	US-09-737-149-29	Sequence 29, Appl
16	46	53.5	3313	16	US-10-038-854-69	Sequence 69, Appl
17	46	53.5	4115	16	US-10-038-854-4	Sequence 4, Appli
18	45	52.3	199	14	US-10-034-934-125	Sequence 125, App
19	45	52.3	434	14	US-10-180-375-124	Sequence 124, App
20	44	51.2	11	14	US-10-161-791-294	Sequence 294, App
21	44	51.2	15	14	US-10-161-791-301	Sequence 301, App
22	44	51.2	86	10	US-09-764-891-2992	Sequence 2992, Ap
23	44	51.2	86	14	US-10-029-386-30668	Sequence 30668, A
24	44	51.2	96	14	US-10-029-386-33742	Sequence 33742, A
25	44	51.2	111	9	US-09-864-761-47005	Sequence 47005, A
26	44	51.2	184	14	US-10-156-761-7948	Sequence 7948, Ap
27	44	51.2	184	14	US-10-029-386-33844	Sequence 33844, A
28	44	51.2	304	11	US-09-833-245-1062	Sequence 1062, Ap
29	44	51.2	304	14	US-10-156-761-13550	Sequence 13550, A
30	44	51.2	350	15	US-10-094-749-1837	Sequence 1837, Ap
31	44	51.2	421	14	US-10-262-666-6	Sequence 6, Appli
32	44	51.2	696	14	US-10-122-805-4	Sequence 4, Appli
33	44	51.2	704	14	US-10-240-154-18	Sequence 18, Appl
34	44	51.2	742	13	US-10-077-111-11	Sequence 11, Appl
35	44	51.2	3338	14	US-10-156-761-8464	Sequence 8464, Ap
36	44	51.2	19695	15	US-10-084-846A-3	Sequence 3, Appli
37	43	50.0	17	9	US-09-938-315-69	Sequence 69, Appl
38	43	50.0	17	14	US-10-161-791-69	Sequence 69, Appl
39	43	50.0	351	13	US-10-004-717-11	Sequence 11, Appl
40	43	50.0	351	13	US-10-004-717-46	Sequence 46, Appl
41	43	50.0	484	9	US-09-738-626-5539	Sequence 5539, Ap
42	43	50.0	688	14	US-10-081-980B-1	Sequence 1, Appli
43	43	50.0	724	9	US-09-962-929-2	Sequence 2, Appli
44	43	50.0	724	9	US-09-962-929-4	Sequence 4, Appli
45	43	50.0	724	14	US-10-081-980B-3	Sequence 3, Appli

ALIGNMENTS

RESULT 1
US-10-181-654-3
; Sequence 3, Application US/10181654
; Publication No. US20030108957A1
; GENERAL INFORMATION:
; APPLICANT: The Wistar Institute of Anatomy and Biology
; APPLICANT: Creighton University
; APPLICANT: Otvos, Laszlo
; APPLICANT: Blaszczyk-Thurin, Magdalena
; APPLICANT: Rogers, Mark
; APPLICANT: Lovas, Sandor
; TITLE OF INVENTION: Biocidal Molecules, Macromolecular Targets and Methods of Produc
; TITLE OF INVENTION: Use
; FILE REFERENCE: WST94BPCT
; CURRENT APPLICATION NUMBER: US/10/181,654
; CURRENT FILING DATE: 2002-07-19
; PRIOR APPLICATION NUMBER: US 60/177,565
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: US 60/237,599
; PRIOR FILING DATE: 2000-10-03
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 3
; LENGTH: 20
; TYPE: PRT
; ORGANISM: P. apterus
US-10-181-654-3

Query Match 94.2%; Score 81; DB 14; Length 20;
Best Local Similarity 87.5%; Pred. No. 0.0013;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy          1 DKGXXLPRPTPPRPIY 16
              ||| |||||||||||
Db          2 DKGSYLPRPTPPRPIY 17
```

RESULT 2

```
US-10-181-654-12
; Sequence 12, Application US/10181654
; Publication No. US20030108957A1
; GENERAL INFORMATION:
;  APPLICANT: The Wistar Institute of Anatomy and Biology
;  APPLICANT:  Creighton University
;  APPLICANT:  Otvos, Laszlo
;  APPLICANT:  Blaszczyk-Thurin, Magdalena
;  APPLICANT:  Rogers, Mark
;  APPLICANT:  Lovas, Sandor
;  TITLE OF INVENTION: Biocidal Molecules, Macromolecular Targets and Methods of Product
;  TITLE OF INVENTION:  Use
;  FILE REFERENCE: WST94BPCT
;  CURRENT APPLICATION NUMBER: US/10/181,654
;  CURRENT FILING DATE:  2002-07-19
;  PRIOR APPLICATION NUMBER: US 60/177,565
;  PRIOR FILING DATE: 2000-01-21
;  PRIOR APPLICATION NUMBER: US 60/237,599
;  PRIOR FILING DATE: 2000-10-03
;  NUMBER OF SEQ ID NOS: 36
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 12
;   LENGTH: 21
;   TYPE: PRT
;   ORGANISM: biotin-K-pyrrhocoricin
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(1)
;   OTHER INFORMATION: biotin is attached to Lys in position 1
US-10-181-654-12

  Query Match             94.2%;  Score 81;  DB 14;  Length 21;
  Best Local Similarity   87.5%;  Pred. No. 0.0014;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 DKGXXLPRPTPPRPIY 16
              |||  ||||||||||||
Db          3 DKGSYLPRPTPPRPIY 18


RESULT 3
US-10-181-654-25
; Sequence 25, Application US/10181654
; Publication No. US20030108957A1
; GENERAL INFORMATION:
;  APPLICANT: The Wistar Institute of Anatomy and Biology
;  APPLICANT:  Creighton University
;  APPLICANT:  Otvos, Laszlo
;  APPLICANT:  Blaszczyk-Thurin, Magdalena
;  APPLICANT:  Rogers, Mark
;  APPLICANT:  Lovas, Sandor
;  TITLE OF INVENTION: Biocidal Molecules, Macromolecular Targets and Methods of Product
;  TITLE OF INVENTION:  Use
;  FILE REFERENCE: WST94BPCT
;  CURRENT APPLICATION NUMBER: US/10/181,654
;  CURRENT FILING DATE:  2002-07-19
;  PRIOR APPLICATION NUMBER: US 60/177,565
;  PRIOR FILING DATE: 2000-01-21
;  PRIOR APPLICATION NUMBER: US 60/237,599
;  PRIOR FILING DATE: 2000-10-03
;  NUMBER OF SEQ ID NOS: 36
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 25
;   LENGTH: 21
;   TYPE: PRT
;   ORGANISM: fluorescein-K pyrrhocoricin
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(1)
;   OTHER INFORMATION: fluorescein is attached to Lys in position 1
US-10-181-654-25

  Query Match             94.2%;  Score 81;  DB 14;  Length 21;
  Best Local Similarity   87.5%;  Pred. No. 0.0014;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 DKGXXLPRPTPPRPIY 16
              |||  ||||||||||||
Db          3 DKGSYLPRPTPPRPIY 18


RESULT 4
US-10-181-654-9
; Sequence 9, Application US/10181654
; Publication No. US20030108957A1
; GENERAL INFORMATION:
;  APPLICANT: The Wistar Institute of Anatomy and Biology
;  APPLICANT:  Creighton University
;  APPLICANT:  Otvos, Laszlo
;  APPLICANT:  Blaszczyk-Thurin, Magdalena
;  APPLICANT:  Rogers, Mark
;  APPLICANT:  Lovas, Sandor
;  TITLE OF INVENTION: Biocidal Molecules, Macromolecular Targets and Methods of Produc
;  TITLE OF INVENTION:  Use
;  FILE REFERENCE: WST94BPCT
;  CURRENT APPLICATION NUMBER: US/10/181,654
;  CURRENT FILING DATE:  2002-07-19
;  PRIOR APPLICATION NUMBER: US 60/177,565
;  PRIOR FILING DATE: 2000-01-21
;  PRIOR APPLICATION NUMBER: US 60/237,599
;  PRIOR FILING DATE: 2000-10-03
;  NUMBER OF SEQ ID NOS: 36
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 9
;   LENGTH: 18
;   TYPE: PRT
;   ORGANISM: modified pyrrhocoricin peptide
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(1)
;   OTHER INFORMATION: A moiety having a net positive charge is attached to Asp
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (4)..(4)
;   OTHER INFORMATION: can be any amino acid
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (5)..(5)
;   OTHER INFORMATION: can be any amino acid
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (17)..(17)
;   OTHER INFORMATION: can be any amino acid
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (18)..(18)
;   OTHER INFORMATION: can be any amino acid and can have attached any of the groups cc
;   OTHER INFORMATION: s
US-10-181-654-9

  Query Match             93.0%;  Score 80;  DB 14;  Length 18;
  Best Local Similarity   100.0%;  Pred. No. 0.0016;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 DKGXXLPRPTPPRPIY 16
              ||||||||||||||||
Db          1 DKGXXLPRPTPPRPIY 16


RESULT 5
US-10-181-654-36
; Sequence 36, Application US/10181654
; Publication No. US20030108957A1
; GENERAL INFORMATION:
;  APPLICANT: The Wistar Institute of Anatomy and Biology
;  APPLICANT:  Creighton University
```

```
;  APPLICANT:  Otvos, Laszlo
;  APPLICANT:  Blaszczyk-Thurin, Magdalena
;  APPLICANT:  Rogers, Mark
;  APPLICANT:  Lovas, Sandor
;  TITLE OF INVENTION: Biocidal Molecules, Macromolecular Targets and Methods of Product
;  TITLE OF INVENTION:  Use
;  FILE REFERENCE: WST94BPCT
;  CURRENT APPLICATION NUMBER: US/10/181,654
;  CURRENT FILING DATE:  2002-07-19
;  PRIOR APPLICATION NUMBER: US 60/177,565
;  PRIOR FILING DATE: 2000-01-21
;  PRIOR APPLICATION NUMBER: US 60/237,599
;  PRIOR FILING DATE: 2000-10-03
;  NUMBER OF SEQ ID NOS: 36
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 36
;   LENGTH: 18
;   TYPE: PRT
;   ORGANISM: modification of Pyrrhocoricin
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(1)
;   OTHER INFORMATION: Asp in position 1 is modified by a 1-aminocyclo-hexane carboxylic
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (18)..(18)
;   OTHER INFORMATION: Arg in position 18 is modified by an amino linker moiety
US-10-181-654-36

  Query Match             86.0%;  Score 74;  DB 14;  Length 18;
  Best Local Similarity   81.2%;  Pred. No. 0.0084;
  Matches   13;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 DKGXXLPRPTPPRPIY 16
              | |  ||||||||||
Db          1 DLGSYLPRPTPPRPIY 16


RESULT 6
US-09-953-342-25
; Sequence 25, Application US/09953342
; Patent No. US20020106735A1
; GENERAL INFORMATION:
;  APPLICANT: Scorilas, Andreas
;  APPLICANT:  Diamandis, Eleftherios
;  TITLE OF INVENTION: NOVEL BCL-2 RELATED PROLINE RICH PROTEIN (BPR)
;  FILE REFERENCE: 11757.52USU1
;  CURRENT APPLICATION NUMBER: US/09/953,342
;  CURRENT FILING DATE:  2001-09-14
;  PRIOR APPLICATION NUMBER: US 60/233,026
;  PRIOR FILING DATE: 2000-09-15
;  NUMBER OF SEQ ID NOS: 30
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 25
;   LENGTH: 176
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-953-342-25

  Query Match             54.7%;  Score 47;  DB 9;  Length 176;
  Best Local Similarity   80.0%;  Pred. No. 1.2e+02;
  Matches    8;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          7 PRPTPPRPIY 16
              | ||||||  |
Db        114 PVPTPPRPSY 123


RESULT 7
US-10-156-761-11324
; Sequence 11324, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
;  APPLICANT: OMURA, SATOSHI
;  APPLICANT:  IKEDA, HARUO
;  APPLICANT:  ISHIKAWA, JUN
;  APPLICANT:  HORIKAWA, HIROSHI
;  APPLICANT:  SHIBA, TADAYOSHI
;  APPLICANT:  SAKAKI, YOSHIYUKI
;  APPLICANT:  HATTORI, MASAHIRA
;  TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
;  FILE REFERENCE: 249-262
;  CURRENT APPLICATION NUMBER: US/10/156,761
;  CURRENT FILING DATE:  2002-05-29
;  PRIOR APPLICATION NUMBER: JP 2001-204089
;  PRIOR FILING DATE: 2001-05-30
;  PRIOR APPLICATION NUMBER: JP 2001-272697
;  PRIOR FILING DATE: 2001-08-02
;  NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 11324
;   LENGTH: 392
;   TYPE: PRT
;   ORGANISM: Streptomyces avermitilis
US-10-156-761-11324

  Query Match             54.7%;  Score 47;  DB 14;  Length 392;
  Best Local Similarity   87.5%;  Pred. No. 2.4e+02;
  Matches    7;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          7 PRPTPPRP 14
              ||||||:|
Db        375 PRPTPPKP 382


RESULT 8
US-10-108-260A-3665
; Sequence 3665, Application US/10108260A
; Publication No. US20040005560A1
; GENERAL INFORMATION:
;  APPLICANT: HELIX RESEARCH INSTITUTE
;  TITLE OF INVENTION: No. US20040005560A1el full length cDNA
;  FILE REFERENCE: H1-A0106
;  CURRENT APPLICATION NUMBER: US/10/108,260A
;  CURRENT FILING DATE:  2002-03-27
;  NUMBER OF SEQ ID NOS: 5458
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3665
;   LENGTH: 1126
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-108-260A-3665

  Query Match             54.7%;  Score 47;  DB 15;  Length 1126;
  Best Local Similarity   77.8%;  Pred. No. 6e+02;
  Matches    7;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          8 RPTPPRPIY 16
              |||||  |:|
Db       1000 RPTPPEPLY 1008


RESULT 9
US-10-181-654-7
; Sequence 7, Application US/10181654
; Publication No. US20030108957A1
; GENERAL INFORMATION:
;  APPLICANT: The Wistar Institute of Anatomy and Biology
;  APPLICANT:  Creighton University
;  APPLICANT:  Otvos, Laszlo
;  APPLICANT:  Blaszczyk-Thurin, Magdalena
;  APPLICANT:  Rogers, Mark
;  APPLICANT:  Lovas, Sandor
;  TITLE OF INVENTION: Biocidal Molecules, Macromolecular Targets and Methods of Produc
;  TITLE OF INVENTION:  Use
;  FILE REFERENCE: WST94BPCT
;  CURRENT APPLICATION NUMBER: US/10/181,654
```

```
;  CURRENT FILING DATE:  2002-07-19
;  PRIOR APPLICATION NUMBER: US 60/177,565
;  PRIOR FILING DATE: 2000-01-21
;  PRIOR APPLICATION NUMBER: US 60/237,599
;  PRIOR FILING DATE: 2000-10-03
;  NUMBER OF SEQ ID NOS: 36
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 7
;   LENGTH: 20
;   TYPE: PRT
;   ORGANISM: insect antibacterial peptide
US-10-181-654-7

  Query Match             53.5%;  Score 46;  DB 14;  Length 20;
  Best Local Similarity   64.3%;  Pred. No. 23;
  Matches    9;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          1 DKGXXLPRPTPPRP 14
              |||   |  || |||
Db          2 DKGRYLEAPTRPRP 15


RESULT 10
US-10-224-999A-3465
; Sequence 3465, Application US/10224999A
; Publication No. US20030171318A1
; GENERAL INFORMATION:
;  APPLICANT: Myriad Genetics, Inc.
;  APPLICANT:  Morham, Scott
;  APPLICANT:  Zavitz, Kenton
;  APPLICANT:  Hobden, Adrian
;  TITLE OF INVENTION: Composition and Method for Treating Viral Infection
;  FILE REFERENCE: 5004.01
;  CURRENT APPLICATION NUMBER: US/10/224,999A
;  CURRENT FILING DATE:  2003-03-03
;  PRIOR APPLICATION NUMBER: US 60/313,695
;  PRIOR FILING DATE: 2001-08-20
;  NUMBER OF SEQ ID NOS: 3484
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3465
;   LENGTH: 487
;   TYPE: PRT
;   ORGANISM: Human herpesvirus 4
US-10-224-999A-3465

  Query Match             53.5%;  Score 46;  DB 14;  Length 487;
  Best Local Similarity   77.8%;  Pred. No. 3.8e+02;
  Matches    7;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          7 PRPTPPRPI 15
              |||||| |:
Db        198 PRPTPPTPL 206


RESULT 11
US-10-149-819-21
; Sequence 21, Application US/10149819
; Publication No. US20030044913A1
; GENERAL INFORMATION:
;  APPLICANT: INCYTE GENOMICS, INC.
;  APPLICANT:  YUE, Henry
;  APPLICANT:  AZIMZAI, Yalda
;  APPLICANT:  TANG, Y. Tom
;  APPLICANT:  PATTERSON, Chandra
;  APPLICANT:  BAUGHN, Mariah R.
;  APPLICANT:  LU, Dyung Aina M.
;  APPLICANT:  SHAH, Purvi
;  APPLICANT:  LAL, Preeti
;  APPLICANT:  AU-YOUNG, Janice
;  APPLICANT:  BURFORD, Neil
;  TITLE OF INVENTION: EXTRACELLULAR MATRIX AND CELL ADHESION MOLECULES
;  FILE REFERENCE: PF-0760 PCT
;  CURRENT APPLICATION NUMBER: US/10/149,819
;  CURRENT FILING DATE:  2002-06-10
;  PRIOR APPLICATION NUMBER: 60/172,852; 60/172,354
;  PRIOR FILING DATE: 1999-12-10; 1999-12-16
;  NUMBER OF SEQ ID NOS: 42
;  SOFTWARE: PERL Program
; SEQ ID NO 21
;   LENGTH: 3298
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: Incyte ID No. US20030044913A1 2847752CD1
US-10-149-819-21

  Query Match             53.5%;  Score 46;  DB 14;  Length 3298;
  Best Local Similarity   61.5%;  Pred. No. 2e+03;
  Matches    8;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 DKGXXLPRPTPPR 13
              |:|  |||  |||
Db       3102 DRGSTLPRRQPPR 3114


RESULT 12
US-10-038-854-68
; Sequence 68, Application US/10038854
; Publication No. US20040022781A1
; GENERAL INFORMATION:
;  APPLICANT: Spytek, Kimberly A
;  APPLICANT:  Li, Li
;  APPLICANT:  Wolenc, Adam R
;  APPLICANT:  Vernet, Corine
;  APPLICANT:  Eisen, Andrew J
;  APPLICANT:  Liu, Xiaohong
;  APPLICANT:  Malyankar, Uriel M
;  APPLICANT:  Shimkets, Richard A
;  APPLICANT:  Tchernev, Velizar
;  APPLICANT:  Spaderna, Steven K
;  APPLICANT:  Gorman, Linda
;  APPLICANT:  Kekuda, Ramesh
;  APPLICANT:  Patturajan, Meera
;  APPLICANT:  Gusev, Vladimir Y
;  APPLICANT:  Gangolli, Esha A
;  APPLICANT:  Guo, Xiaojia S
;  APPLICANT:  Shenoy, Suresh G
;  APPLICANT:  Rastelli, Luca
;  APPLICANT:  Casman, Stacie J
;  APPLICANT:  Boldog, Ferenc
;  APPLICANT:  Burgess, Catherine E
;  APPLICANT:  Edinger, Shlomit R
;  APPLICANT:  Ellerman, Karen
;  APPLICANT:  Gunther, Erik
;  APPLICANT:  Smithson, Glennda
;  APPLICANT:  Millet, Isabelle
;  APPLICANT:  MacDougall, John R
;  TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
;  FILE REFERENCE: 21402-230
;  CURRENT APPLICATION NUMBER: US/10/038,854
;  CURRENT FILING DATE:  2003-01-22
;  PRIOR APPLICATION NUMBER: 60/258,928
;  PRIOR FILING DATE: 2000-12-29
;  PRIOR APPLICATION NUMBER: 60/259,415
;  PRIOR FILING DATE: 2001-01-02
;  PRIOR APPLICATION NUMBER: 60/259,785
;  PRIOR FILING DATE: 2001-01-04
;  PRIOR APPLICATION NUMBER: 60/269,814
;  PRIOR FILING DATE: 2001-02-20
;  PRIOR APPLICATION NUMBER: 60/279,832
;  PRIOR FILING DATE: 2001-03-29
;  PRIOR APPLICATION NUMBER: 60/279,833
;  PRIOR FILING DATE: 2001-03-29
;  PRIOR APPLICATION NUMBER: 60/279,863
;  PRIOR FILING DATE: 2001-03-29
```

```
;  PRIOR APPLICATION NUMBER: 60/283,889
;  PRIOR FILING DATE: 2001-04-13
;  PRIOR APPLICATION NUMBER: 60/284,447
;  PRIOR FILING DATE: 2001-04-18
;  PRIOR APPLICATION NUMBER: 60/286,683
;  PRIOR FILING DATE: 2001-04-25
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 411
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 68
;   LENGTH: 3301
;   TYPE: PRT
;   ORGANISM: Mus musculus
US-10-038-854-68

  Query Match           53.5%;  Score 46;  DB 16;  Length 3301;
  Best Local Similarity   61.5%;  Pred. No. 2e+03;
  Matches    8;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 DKGXXLPRPTPPR 13
              |:|  |||  |||
Db       3109 DRGSTLPRRQPPR 3121


RESULT 13
US-10-225-567A-656
; Sequence 656, Application US/10225567A
; Publication No. US20030113798A1
; GENERAL INFORMATION:
;  APPLICANT: LifeSpan Biosciences
;  APPLICANT:  Brown, Joseph P.
;  APPLICANT:  Burmer, Glenna C.
;  APPLICANT:  Roush, Christine L.
;  TITLE OF INVENTION: ANTIGENIC PEPTIDES AND ANTIBODIES FOR G PROTEIN-COUPLED RECEPTORS
;  FILE REFERENCE: 1920-4-4
;  CURRENT APPLICATION NUMBER: US/10/225,567A
;  CURRENT FILING DATE:  2001-12-19
;  PRIOR APPLICATION NUMBER: 60/257,144
;  PRIOR FILING DATE: 2000-12-19
;  NUMBER OF SEQ ID NOS: 2292
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 656
;   LENGTH: 3312
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-225-567A-656

  Query Match           53.5%;  Score 46;  DB 14;  Length 3312;
  Best Local Similarity   61.5%;  Pred. No. 2e+03;
  Matches    8;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 DKGXXLPRPTPPR 13
              |:|  |||  |||
Db       3116 DRGSTLPRRQPPR 3128


RESULT 14
US-10-038-854-67
; Sequence 67, Application US/10038854
; Publication No. US20040022781A1
; GENERAL INFORMATION:
;  APPLICANT: Spytek, Kimberly A
;  APPLICANT:  Li, Li
;  APPLICANT:  Wolenc, Adam R
;  APPLICANT:  Vernet, Corine
;  APPLICANT:  Eisen, Andrew J
;  APPLICANT:  Liu, Xiaohong
;  APPLICANT:  Malyankar, Uriel M
;  APPLICANT:  Shimkets, Richard A
;  APPLICANT:  Tchernev, Velizar
;  APPLICANT:  Spaderna, Steven K
;  APPLICANT:  Gorman, Linda
;  APPLICANT:  Kekuda, Ramesh
;  APPLICANT:  Patturajan, Meera
;  APPLICANT:  Gusev, Vladimir Y
;  APPLICANT:  Gangolli, Esha A
;  APPLICANT:  Guo, Xiaojia S
;  APPLICANT:  Shenoy, Suresh G
;  APPLICANT:  Rastelli, Luca
;  APPLICANT:  Casman, Stacie J
;  APPLICANT:  Boldog, Ferenc
;  APPLICANT:  Burgess, Catherine E
;  APPLICANT:  Edinger, Shlomit R
;  APPLICANT:  Ellerman, Karen
;  APPLICANT:  Gunther, Erik
;  APPLICANT:  Smithson, Glennda
;  APPLICANT:  Millet, Isabelle
;  APPLICANT:  MacDougall, John R
;  TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
;  FILE REFERENCE: 21402-230
;  CURRENT APPLICATION NUMBER: US/10/038,854
;  CURRENT FILING DATE:  2003-01-22
;  PRIOR APPLICATION NUMBER: 60/258,928
;  PRIOR FILING DATE: 2000-12-29
;  PRIOR APPLICATION NUMBER: 60/259,415
;  PRIOR FILING DATE: 2001-01-02
;  PRIOR APPLICATION NUMBER: 60/259,785
;  PRIOR FILING DATE: 2001-01-04
;  PRIOR APPLICATION NUMBER: 60/269,814
;  PRIOR FILING DATE: 2001-02-20
;  PRIOR APPLICATION NUMBER: 60/279,832
;  PRIOR FILING DATE: 2001-03-29
;  PRIOR APPLICATION NUMBER: 60/279,833
;  PRIOR FILING DATE: 2001-03-29
;  PRIOR APPLICATION NUMBER: 60/279,863
;  PRIOR FILING DATE: 2001-03-29
;  PRIOR APPLICATION NUMBER: 60/283,889
;  PRIOR FILING DATE: 2001-04-13
;  PRIOR APPLICATION NUMBER: 60/284,447
;  PRIOR FILING DATE: 2001-04-18
;  PRIOR APPLICATION NUMBER: 60/286,683
;  PRIOR FILING DATE: 2001-04-25
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 411
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 67
;   LENGTH: 3312
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-038-854-67

  Query Match           53.5%;  Score 46;  DB 16;  Length 3312;
  Best Local Similarity   61.5%;  Pred. No. 2e+03;
  Matches    8;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 DKGXXLPRPTPPR 13
              |:|  |||  |||
Db       3116 DRGSTLPRRQPPR 3128


RESULT 15
US-09-737-149-29
; Sequence 29, Application US/09737149
; Patent No. US20020077466A1
; GENERAL INFORMATION:
;  APPLICANT: Spaderna, Steven K
;  APPLICANT:  Quinn, Kerry E.
;  APPLICANT:  Shimkets, Richard A.
;  APPLICANT:  Muralidhara, Padigaru
;  APPLICANT:  Spytek, Kimberly A.
;  TITLE OF INVENTION: Polypeptides and Nucleic Acids Encoding Same
;  FILE REFERENCE: 15966-620 CIP
;  CURRENT APPLICATION NUMBER: US/09/737,149
;  CURRENT FILING DATE:  2001-06-15
;  PRIOR APPLICATION NUMBER: 60/170,564
;  PRIOR FILING DATE: 1999-12-14
```

```
;  PRIOR APPLICATION NUMBER: 60/173,165
;  PRIOR FILING DATE: 1999-12-27
;  PRIOR APPLICATION NUMBER: 60/173,362
;  PRIOR FILING DATE: 1999-12-27
;  PRIOR APPLICATION NUMBER: 60/173,544
;  PRIOR FILING DATE: 1999-12-29
;  PRIOR APPLICATION NUMBER: 60/174,404
;  PRIOR FILING DATE: 2000-01-04
;  PRIOR APPLICATION NUMBER: 60/174,962
;  PRIOR FILING DATE: 2000-01-07
;  PRIOR APPLICATION NUMBER: 60/223,929
;  PRIOR FILING DATE: 2000-08-09
;  NUMBER OF SEQ ID NOS: 49
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 29
;   LENGTH: 3313
;   TYPE: PRT
;   ORGANISM: Rattus norvegicus
US-09-737-149-29

  Query Match             53.5%;  Score 46;  DB 9;  Length 3313;
  Best Local Similarity   61.5%;  Pred. No. 2e+03;
  Matches    8;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 DKGXXLPRPTPPR 13
              |:|  |||  |||
Db       3117 DRGSTLPRRQPPR 3129


Search completed: March 11, 2004, 17:01:06
Job time : 35 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: March 11, 2004, 16:53:29 ; Search time 23 Seconds
(without alignments)
40.403 Million cell updates/sec

Title: US-09-980-804-1
Perfect score: 86
Sequence: 1 DKGXXLPRPTPPRPIYXX 18

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	50	58.1	167	4	US-09-252-991A-23665	Sequence 23665, A
2	47	54.7	361	4	US-09-252-991A-28125	Sequence 28125, A
3	46	53.5	884	2	US-08-465-976A-2	Sequence 2, Appli
4	46	53.5	884	2	US-08-982-412-2	Sequence 2, Appli
5	46	53.5	902	1	US-08-396-479B-6	Sequence 6, Appli
6	46	53.5	902	1	US-08-818-823-6	Sequence 6, Appli
7	45	52.3	412	4	US-09-252-991A-26237	Sequence 26237, A
8	45	52.3	1255	2	US-09-080-897-4	Sequence 4, Appli
9	45	52.3	1255	3	US-08-899-595-1	Sequence 1, Appli
10	45	52.3	1255	3	US-09-323-735-4	Sequence 4, Appli
11	44	51.2	11	3	US-08-602-999A-294	Sequence 294, App
12	44	51.2	11	3	US-08-652-877-36	Sequence 36, Appl
13	44	51.2	11	3	US-08-476-515A-36	Sequence 36, Appl
14	44	51.2	11	4	US-09-500-124-294	Sequence 294, App
15	44	51.2	15	3	US-08-602-999A-301	Sequence 301, App
16	44	51.2	15	4	US-09-500-124-301	Sequence 301, App
17	44	51.2	156	4	US-09-732-210-1643	Sequence 1643, Ap
18	44	51.2	256	4	US-09-252-991A-25943	Sequence 25943, A
19	44	51.2	330	4	US-09-252-991A-18388	Sequence 18388, A
20	44	51.2	418	4	US-09-252-991A-17463	Sequence 17463, A
21	44	51.2	486	4	US-09-252-991A-31404	Sequence 31404, A
22	44	51.2	517	4	US-09-252-991A-32085	Sequence 32085, A
23	44	51.2	696	3	US-08-906-865-4	Sequence 4, Appli
24	44	51.2	696	4	US-09-129-668-4	Sequence 4, Appli
25	43	50.0	17	3	US-08-602-999A-69	Sequence 69, Appl
26	43	50.0	17	4	US-08-278-865-69	Sequence 69, Appl
27	43	50.0	17	4	US-09-500-124-69	Sequence 69, Appl
28	43	50.0	108	4	US-09-489-039A-12459	Sequence 12459, A
29	43	50.0	664	4	US-09-252-991A-31116	Sequence 31116, A
30	43	50.0	722	3	US-08-390-874C-12	Sequence 12, Appl
31	43	50.0	722	4	US-09-265-772-12	Sequence 12, Appl
32	43	50.0	724	1	US-07-906-349A-5	Sequence 5, Appli
33	43	50.0	724	1	US-08-167-035-2	Sequence 2, Appli
34	43	50.0	724	1	US-08-208-887A-2	Sequence 2, Appli
35	43	50.0	724	2	US-08-539-005-2	Sequence 2, Appli
36	43	50.0	724	4	US-09-280-598-5	Sequence 5, Appli
37	43	50.0	724	4	US-09-963-137-179	Sequence 179, App
38	43	50.0	724	4	US-09-963-137-181	Sequence 181, App
39	43	50.0	968	4	US-09-417-197-49	Sequence 49, Appl
40	43	50.0	970	4	US-09-417-197-67	Sequence 67, Appl
41	42.5	49.4	791	4	US-09-252-991A-27312	Sequence 27312, A
42	42	48.8	105	4	US-09-288-143-99	Sequence 99, Appl
43	42	48.8	393	3	US-08-888-429A-21	Sequence 21, Appl
44	42	48.8	393	4	US-09-593-653-21	Sequence 21, Appl
45	42	48.8	487	4	US-09-206-166-6	Sequence 6, Appli

ALIGNMENTS

RESULT 1
US-09-252-991A-23665
; Sequence 23665, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 23665
; LENGTH: 167
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-23665

Query Match 58.1%; Score 50; DB 4; Length 167;
Best Local Similarity 100.0%; Pred. No. 5.2;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          7 PRPTPPRP 14
              ||||||||
Db         93 PRPTPPRP 100
```

RESULT 2
US-09-252-991A-28125
; Sequence 28125, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 28125
; LENGTH: 361
; TYPE: PRT

```
;   ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-28125

  Query Match             54.7%;  Score 47;  DB 4;  Length 361;
  Best Local Similarity   61.5%;  Pred. No. 29;
  Matches    8;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          2 KGXXLPRPTPPRP 14
              :|  |||| | ||
Db        266 RGPALPRPAPARP 278


RESULT 3
US-08-465-976A-2
; Sequence 2, Application US/08465976A
; Patent No. 5869632
;  GENERAL INFORMATION:
;    APPLICANT:  SOPPET, DANIEL R
;    APPLICANT:  LI, YI
;    APPLICANT:  ROSEN, CRAIG A
;    APPLICANT:  RUBEN, STEVEN M
;    TITLE OF INVENTION:  HUMAN G-PROTEIN RECEPTOR
;    NUMBER OF SEQUENCES:  7
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  CARELLA, BYRNE, BAIN GILFILLAN, CECCHI
;      ADDRESSEE:  STEWART & OLSTEIN
;      STREET:  6 BECKER FARM ROAD
;      CITY:  ROSELAND
;      STATE:  NJ
;      COUNTRY:  US
;      ZIP:  07068
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/465,976A
;      FILING DATE:  06-JUN-1995
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  FERRARO, GREGORY F
;      REGISTRATION NUMBER:  36,134
;      REFERENCE/DOCKET NUMBER:  325800-444
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (201) 994-1700
;      TELEFAX:  (201) 994-1744
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  884 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-465-976A-2

  Query Match             53.5%;  Score 46;  DB 2;  Length 884;
  Best Local Similarity   61.5%;  Pred. No. 95;
  Matches    8;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 DKGXXLPRPTPPR 13
              |:|  |||  |||
Db        678 DRGSTLPRRQPPR 690


RESULT 4
US-08-982-412-2
; Sequence 2, Application US/08982412
; Patent No. 5958729
;  GENERAL INFORMATION:
;    APPLICANT:  SOPPET, DANIEL R
;    APPLICANT:  LI, YI
;    APPLICANT:  ROSEN, CRAIG A
;    APPLICANT:  RUBEN, STEVEN M
;    TITLE OF INVENTION:  HUMAN G-PROTEIN RECEPTOR
;    NUMBER OF SEQUENCES:  7
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  HUMAN GENOME SCIENCES, INC.
;      STREET:  9410 KEY WEST AVENUE
;      CITY:  ROCKVILLE,
;      STATE:  MD
;      COUNTRY:  US
;      ZIP:  20850
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/982,412
;      FILING DATE:
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  BROOKES, ANDERS A
;      REGISTRATION NUMBER:  36,373
;      REFERENCE/DOCKET NUMBER:  PF181PCT2
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (301) 309-8504
;      TELEFAX:  (301) 309-8439
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  884 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-982-412-2

  Query Match             53.5%;  Score 46;  DB 2;  Length 884;
  Best Local Similarity   61.5%;  Pred. No. 95;
  Matches    8;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 DKGXXLPRPTPPR 13
              |:|  |||  |||
Db        678 DRGSTLPRRQPPR 690


RESULT 5
US-08-396-479B-6
; Sequence 6, Application US/08396479B
; Patent No. 5612455
;  GENERAL INFORMATION:
;    APPLICANT:  HOEY, Timothy
;    TITLE OF INVENTION:  NUCLEAR FACTORS AND BINDING ASSAY
;    NUMBER OF SEQUENCES:  18
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  FLEHR, HOHBACH, TEST, ALBRITTON & HERBERT
;      STREET:  4 Embarcadero Center, Suite 3400
;      CITY:  San Francisco
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94111
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/396,479B
;      FILING DATE:
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Osman, Richard A
;      REGISTRATION NUMBER:  36,627
;      REFERENCE/DOCKET NUMBER:  A-59450-1/RAO
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 494-8700
;      TELEFAX:  (415) 494-8771
```

```
;      TELEX:  210 277299
;  INFORMATION FOR SEQ ID NO:  6:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  902 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-396-479B-6

  Query Match             53.5%;  Score 46;  DB 1;  Length 902;
  Best Local Similarity   77.8%;  Pred. No. 97;
  Matches    7;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          6 LPRPTPPRP 14
              :||| ||||
Db         61 IPRPPPPRP 69


RESULT 6
US-08-818-823-6
; Sequence 6, Application US/08818823
; Patent No. 5708158
;  GENERAL INFORMATION:
;    APPLICANT:  HOEY, Timothy
;    TITLE OF INVENTION:  NUCLEAR FACTORS AND BINDING ASSAY
;    NUMBER OF SEQUENCES:  18
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  FLEHR, HOHBACH, TEST, ALBRITTON & HERBERT
;      STREET:  4 Embarcadero Center, Suite 3400
;      CITY:  San Francisco
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94111
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/818,823
;      FILING DATE:  14-MAR-1997
;      CLASSIFICATION:  536
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/396,479
;      FILING DATE:  02-MAR-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Osman, Richard A
;      REGISTRATION NUMBER:  36,627
;      REFERENCE/DOCKET NUMBER:  A-59450-1/RAO
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 494-8700
;      TELEFAX:  (415) 494-8771
;      TELEX:  210 277299
;  INFORMATION FOR SEQ ID NO:  6:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  902 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-818-823-6

  Query Match             53.5%;  Score 46;  DB 1;  Length 902;
  Best Local Similarity   77.8%;  Pred. No. 97;
  Matches    7;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          6 LPRPTPPRP 14
              :||| ||||
Db         61 IPRPPPPRP 69


RESULT 7
US-09-252-991A-26237
; Sequence 26237, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
;  APPLICANT:  Marc J. Rubenfield et al.
;  TITLE OF INVENTION:  NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
;  TITLE OF INVENTION:  AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE:  107196.136
;  CURRENT APPLICATION NUMBER: US/09/252,991A
;  CURRENT FILING DATE:  1999-02-18
;  PRIOR APPLICATION NUMBER:  US 60/074,788
;  PRIOR FILING DATE:  1998-02-18
;  PRIOR APPLICATION NUMBER:  US 60/094,190
;  PRIOR FILING DATE:  1998-07-27
;  NUMBER OF SEQ ID NOS:  33142
; SEQ ID NO 26237
;   LENGTH: 412
;   TYPE: PRT
;   ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-26237

  Query Match             52.3%;  Score 45;  DB 4;  Length 412;
  Best Local Similarity   87.5%;  Pred. No. 61;
  Matches    7;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          7 PRPTPPRP 14
              ||| ||||
Db        400 PRPAPPRP 407


RESULT 8
US-09-080-897-4
; Sequence 4, Application US/09080897
; Patent No. 5985574
;  GENERAL INFORMATION:
;    APPLICANT:  King, Mary-Claire
;    APPLICANT:  Lynch, Eric D.
;    APPLICANT:  Lee, Ming
;    APPLICANT:  Morrow, Jan E.
;    APPLICANT:  Welcsh, Piri L.
;    APPLICANT:  Leon, Pedro E.
;    TITLE OF INVENTION:  Modulators of Actin
;    NUMBER OF SEQUENCES:  14
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  SCIENCE & TECHNOLOGY LAW GROUP
;      STREET:  75 DENISE DRIVE
;      CITY:  HILLSBOROUGH
;      STATE:  CALIFORNIA
;      COUNTRY:  USA
;      ZIP:  94010
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/080,897
;      FILING DATE:
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  OSMAN, RICHARD A
;      REGISTRATION NUMBER:  36,627
;      REFERENCE/DOCKET NUMBER:  UW97-001
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (650) 343-4341
;      TELEFAX:  (650) 343-4342
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1255 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-09-080-897-4
```

```
  Query Match             52.3%;  Score 45;  DB 2;  Length 1255;
  Best Local Similarity   46.7%;  Pred. No. 1.8e+02;
  Matches    7;  Conservative    2;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 DKGXXLPRPTPPRPI 15
              | |  :| | || |:
Db        580 DSGTVIPPPPPPPPL 594


RESULT 9
US-08-899-595-1
; Sequence 1, Application US/08899595
; Patent No. 6111072
;  GENERAL INFORMATION:
;    APPLICANT:  Narumiya, Shuh
;    APPLICANT:  Takahashi, No. 6111072uaki
;    TITLE OF INVENTION:  RHO TARGET PROTEIN HUMAN MDIA AND GENE
;    TITLE OF INVENTION:  ENCODING SAME
;    NUMBER OF SEQUENCES:  14
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Foley & Lardner
;      STREET:  3000 K Street, N.W., Suite 500
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  USA
;      ZIP:  20007-5109
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/899,595
;      FILING DATE:  24-JUL-1997
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 8-242701
;      FILING DATE:  26-AUG-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 9-90170
;      FILING DATE:  25-MAR-1997
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Stephen A. Bent
;      REGISTRATION NUMBER:  29,768
;      REFERENCE/DOCKET NUMBER:  049441/0112
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202)672-5300
;      TELEFAX:  (202)672-5399
;      TELEX:  904136
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1255 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-899-595-1

  Query Match             52.3%;  Score 45;  DB 3;  Length 1255;
  Best Local Similarity   46.7%;  Pred. No. 1.8e+02;
  Matches    7;  Conservative    2;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 DKGXXLPRPTPPRPI 15
              | |  :| | || |:
Db        580 DSGTVIPPPPPPPPL 594


RESULT 10
US-09-323-735-4
; Sequence 4, Application US/09323735
; Patent No. 6197932
;  GENERAL INFORMATION:
;    APPLICANT:  King, Mary-Claire
;    APPLICANT:  Lynch, Eric D.
;    APPLICANT:  Lee, Ming
;    APPLICANT:  Morrow, Jan E.
;    APPLICANT:  Welcsh, Piri L.
;    APPLICANT:  Leon, Pedro E.
;    TITLE OF INVENTION:  Modulators of Actin
;    NUMBER OF SEQUENCES:  14
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  SCIENCE & TECHNOLOGY LAW GROUP
;      STREET:  75 DENISE DRIVE
;      CITY:  HILLSBOROUGH
;      STATE:  CALIFORNIA
;      COUNTRY:  USA
;      ZIP:  94010
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/323,735
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  09/080,897
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  OSMAN, RICHARD A
;      REGISTRATION NUMBER:  36,627
;      REFERENCE/DOCKET NUMBER:  UW97-001
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (650) 343-4341
;      TELEFAX:  (650) 343-4342
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1255 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-09-323-735-4

  Query Match             52.3%;  Score 45;  DB 3;  Length 1255;
  Best Local Similarity   46.7%;  Pred. No. 1.8e+02;
  Matches    7;  Conservative    2;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 DKGXXLPRPTPPRPI 15
              | |  :| | || |:
Db        580 DSGTVIPPPPPPPPL 594


RESULT 11
US-08-602-999A-294
; Sequence 294, Application US/08602999A
; Patent No. 6184205
;  GENERAL INFORMATION:
;    APPLICANT:  SPARKS, Andrew B.
;    APPLICANT:  KAY, Brian K.
;    APPLICANT:  THORN, Judith M.
;    APPLICANT:  QUILLIAM, Lawrence A.
;    APPLICANT:  DER, Channing J.
;    APPLICANT:  FOWLKES, Dana M.
;    APPLICANT:  RIDER, James E.
;    TITLE OF INVENTION:  SH3 BINDING PEPTIDES AND METHODS OF
;    TITLE OF INVENTION:  ISOLATING AND USING SAME
;    NUMBER OF SEQUENCES:  467
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Pennie & Edmonds
;      STREET:  1155 Avenue of the Americas
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  U.S.A.
;      ZIP:  10036-2711
;    COMPUTER READABLE FORM:
```

```
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.30
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/602,999A
;       FILING DATE:  16-FEB-1996
;       CLASSIFICATION:  435
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Misrock, S. Leslie
;       REGISTRATION NUMBER:  18,872
;       REFERENCE/DOCKET NUMBER:  1101-202
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (212) 790-9090
;       TELEFAX:  (212) 869-9741/8864
;       TELEX:  66141 PENNIE
;   INFORMATION FOR SEQ ID NO:  294:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  11 amino acids
;       TYPE:  amino acid
;       TOPOLOGY:  unknown
;     MOLECULE TYPE:  peptide
US-08-602-999A-294

  Query Match             51.2%;  Score 44;  DB 3;  Length 11;
  Best Local Similarity   77.8%;  Pred. No. 2.3;
  Matches    7;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          6 LPRPTPPRP 14
              || | ||||
Db          2 LPEPAPPRP 10


RESULT 12
US-08-652-877-36
; Sequence 36, Application US/08652877
; Patent No. 6187548
;  GENERAL INFORMATION:
;     APPLICANT:  Akerstrom, Goran
;     APPLICANT:  Juhlin, Claes
;     APPLICANT:  Rask, Lars
;     APPLICANT:  Crumley, Gregg R.
;     APPLICANT:  Morse, Clarence C.
;     APPLICANT:  Murray, Edward M.
;     APPLICANT:  Hjalm, Goran
;     TITLE OF INVENTION:  Human Calcium Sensor Protein, Fragments
;     TITLE OF INVENTION:  Thereof and DNA Encoding Same
;     NUMBER OF SEQUENCES:  106
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Rhone-Poulenc Rorer Inc.
;       STREET:  500 Arcola Rd., 3C43
;       CITY:  Collegeville
;       STATE:  PA
;       COUNTRY:  USA
;       ZIP:  19426-0107
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  Macintosh
;       OPERATING SYSTEM:  System 7.5.1
;       SOFTWARE:  Word 6.0 (Patentin)
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/652,877
;       FILING DATE:
;       CLASSIFICATION:  435
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  PCT/US95/15203
;       FILING DATE:  22-NOV-1995
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/344,836
;       FILING DATE:  23-NOV-1994
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/487,314
;       FILING DATE:  07-JUNE-1995
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Savitzky, Martin
;       REGISTRATION NUMBER:  29,699
;       REFERENCE/DOCKET NUMBER:  A1355E-US
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  610-454-3816
;       TELEFAX:  610-454-3808
;   INFORMATION FOR SEQ ID NO:  36:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  11 amino acids
;       TYPE:  amino acid
;       STRANDEDNESS:
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  peptide
;     HYPOTHETICAL:  NO
;     FRAGMENT TYPE:  internal
US-08-652-877-36

  Query Match             51.2%;  Score 44;  DB 3;  Length 11;
  Best Local Similarity   87.5%;  Pred. No. 2.3;
  Matches    7;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          7 PRPTPPRP 14
              ||| ||||
Db          2 PRPLPPRP 9


RESULT 13
US-08-476-515A-36
; Sequence 36, Application US/08476515A
; Patent No. 6239270
;  GENERAL INFORMATION:
;     APPLICANT:  Akerstrom, Goran
;     APPLICANT:  Juhlin, Claes
;     APPLICANT:  Rask, Lars
;     APPLICANT:  Crumley, Gregg R.
;     APPLICANT:  Morse, Clarence C.
;     APPLICANT:  Murray, Edward M.
;     APPLICANT:  Hjalm, Goran
;     TITLE OF INVENTION:  Human Calcium Sensor Protein, Fragments
;     TITLE OF INVENTION:  Thereof and DNA Encoding Same
;     NUMBER OF SEQUENCES:  84
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Martin Savitzky
;       STREET:  Rhone-Poulenc Rorer Inc., 500 Arcola Rd.;
;       STREET:  3C43,
;       CITY:  Collegeville
;       STATE:  PA
;       COUNTRY:  USA
;       ZIP:  19426-0107
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  Compaq PC
;       OPERATING SYSTEM:  Windows 95
;       SOFTWARE:  Word 7.0 (Patentin)
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/476,515A
;       FILING DATE:  07-JUN-1995
;       CLASSIFICATION:  435
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/344,836
;       FILING DATE:  23-NOV-1994
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  WO PCT/SE94/00483
;       FILING DATE:  24-MAY-1994
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  SE 9301764-8
;       FILING DATE:  24-MAY-1993
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Savitzky, Martin
;       REGISTRATION NUMBER:  29,699
;       REFERENCE/DOCKET NUMBER:  A1355D
;     TELECOMMUNICATION INFORMATION:
```

```
;      TELEPHONE:  610-454-3816
;      TELEFAX:  610-454-3808
;  INFORMATION FOR SEQ ID NO:  36:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  11 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    HYPOTHETICAL:  NO
;    FRAGMENT TYPE:  internal
US-08-476-515A-36

  Query Match             51.2%;  Score 44;  DB 3;  Length 11;
  Best Local Similarity   87.5%;  Pred. No. 2.3;
  Matches    7;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          7 PRPTPPRP 14
              ||| ||||
Db          2 PRPLPPRP 9


RESULT 14
US-09-500-124-294
; Sequence 294, Application US/09500124
; Patent No. 6432920
;  GENERAL INFORMATION:
;    APPLICANT:  SPARKS, Andrew B.
;    APPLICANT:  KAY, Brian K.
;    APPLICANT:  THORN, Judith M.
;    APPLICANT:  QUILLIAM, Lawrence A.
;    APPLICANT:  DER, Channing J.
;    APPLICANT:  FOWLKES, Dana M.
;    APPLICANT:  RIDER, James E.
;    TITLE OF INVENTION:  SH3 BINDING PEPTIDES AND METHODS OF
;    TITLE OF INVENTION:  ISOLATING AND USING SAME
;    NUMBER OF SEQUENCES:  467
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Pennie & Edmonds
;      STREET:  1155 Avenue of the Americas
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  U.S.A.
;      ZIP:  10036-2711
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/500,124
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/602,999
;      FILING DATE:  16-FEB-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Misrock, S. Leslie
;      REGISTRATION NUMBER:  18,872
;      REFERENCE/DOCKET NUMBER:  1101-202
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 790-9090
;      TELEFAX:  (212) 869-9741/8864
;      TELEX:  66141 PENNIE
;  INFORMATION FOR SEQ ID NO:  294:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  11 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  unknown
;    MOLECULE TYPE:  peptide
US-09-500-124-294

  Query Match             51.2%;  Score 44;  DB 4;  Length 11;
  Best Local Similarity   77.8%;  Pred. No. 2.3;
  Matches    7;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          6 LPRPTPPRP 14
              || | ||||
Db          2 LPEPAPPRP 10


RESULT 15
US-08-602-999A-301
; Sequence 301, Application US/08602999A
; Patent No. 6184205
;  GENERAL INFORMATION:
;    APPLICANT:  SPARKS, Andrew B.
;    APPLICANT:  KAY, Brian K.
;    APPLICANT:  THORN, Judith M.
;    APPLICANT:  QUILLIAM, Lawrence A.
;    APPLICANT:  DER, Channing J.
;    APPLICANT:  FOWLKES, Dana M.
;    APPLICANT:  RIDER, James E.
;    TITLE OF INVENTION:  SH3 BINDING PEPTIDES AND METHODS OF
;    TITLE OF INVENTION:  ISOLATING AND USING SAME
;    NUMBER OF SEQUENCES:  467
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Pennie & Edmonds
;      STREET:  1155 Avenue of the Americas
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  U.S.A.
;      ZIP:  10036-2711
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/602,999A
;      FILING DATE:  16-FEB-1996
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Misrock, S. Leslie
;      REGISTRATION NUMBER:  18,872
;      REFERENCE/DOCKET NUMBER:  1101-202
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 790-9090
;      TELEFAX:  (212) 869-9741/8864
;      TELEX:  66141 PENNIE
;  INFORMATION FOR SEQ ID NO:  301:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  15 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  unknown
;    MOLECULE TYPE:  peptide
US-08-602-999A-301

  Query Match             51.2%;  Score 44;  DB 3;  Length 15;
  Best Local Similarity   87.5%;  Pred. No. 3.1;
  Matches    7;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          7 PRPTPPRP 14
              ||| ||||
Db          4 PRPLPPRP 11


Search completed: March 11, 2004, 16:56:41
Job time : 24 secs
```

```
!!AA_SEQUENCE 1.0
ID   AAR50300 standard; peptide; 20 AA.
XX
AC   AAR50300;
XX
DT   25-MAR-2003  (revised)
DT   10-OCT-1994  (first entry)
XX
DE   Anti-bacterial glycopeptide #9 induced in Pyrrhocoris apterus.
XX
KW   Antibacterial glycopeptide; Diptera; septicaemia; Gram positive bacteria;
KW   Gram negative bacteria.
XX
OS   Pyrrhocoris apterus.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   11
FT                   /label= O-glycosylated
XX
PN   WO9405787-A1.
XX
PD   17-MAR-1994.
XX
PF   06-SEP-1993;   93WO-FR000853.
XX
PR   04-SEP-1992;   92FR-00010608.
XX
PA   (CNRS ) CNRS CENT NAT RECH SCI.
XX
PI   Bulet P,  Hetru C,  Dimarcq J,  Hoffmann J,  Van Dorsselaer A;
XX
DR   WPI; 1994-101192/12.
XX
PT   New antibacterial glycopeptide(s) derived from insects - for control of
PT   Gram negative and positive bacteria in human and veterinary medicine,
PT   agriculture, etc.
XX
PS   Claim 17; Page 9-10; 45pp; French.
XX
CC   This is a preferred example of an anti-bacterial glycopeptide induced in
CC   arthropods (esp. larval or adult insects) by injection of bacteria, a
CC   septic wound or other injury. The peptides contain at least one O-
CC   glycosylated residue and are useful for treatment of e.g. septicaemia,
CC   for oral or dental use and in gynaecology. (Updated on 25-MAR-2003 to
CC   correct PN field.)
XX
SQ   Sequence 20 AA;

  AAR50300  Length: 20  March 11, 2004 17:24  Type: P  Check: 6865  ..

       1  VDKGSYLPRP TPPRPIYNRN
```

```
!!AA_SEQUENCE 1.0
ID   AAG62740 standard; peptide; 18 AA.
XX
AC   AAG62740;
XX
DT   17-SEP-2001  (first entry)
XX
DE   Amino acid sequence of modified antibacterial peptide pyrrhocoricin.
XX
KW   Multi-helical lid; heat shock protein; hsp; protein folding;
KW   pathogenic infection; bacterial infection; antibacterial.
XX
OS   Unidentified.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   1
FT                   /note= "a moiety having a net positive charge is
FT                   attached"
XX
PN   WO200153509-A2.
XX
PD   26-JUL-2001.
XX
PF   19-JAN-2001; 2001WO-US001812.
XX
PR   21-JAN-2000; 2000US-0177565P.
PR   03-OCT-2000; 2000US-0237599P.
XX
PA   (WIST-) WISTAR INST ANATOMY & BIOLOGY.
PA   (UYCR-) UNIV CREIGHTON.
XX
PI   Otvos L,  Blaszczyk-Thurin M,  Rogers M,  Lovas S;
XX
DR   WPI; 2001-451911/48.
XX
PT   Composition, used to treat a pathogenic infection and eliminate a plant,
PT   insect, or animal pest, comprises a molecule that binds to a heat shock
PT   protein.
XX
PS   Disclosure; Page 111; 124pp; English.
XX
CC   The specification describes a composition that comprises a synthetic non-
CC   naturally occurring molecule that binds to a selected multi-helical lid
CC   of a heat shock protein (hsp) of a selected organism, where the molecule
CC   inhibits protein folding activity of the hsp, and a carrier, where
CC   exposure of the organism to the composition retards the growth and
CC   reproduction of the organism. The composition is used to treat a mammal
CC   suffering from a pathogenic infection, in the manufacture of a medicament
CC   for treating a mammal for a pathogenic infection, and to eliminate a
CC   plant, insect, or animal pest. It is used in the manufacture of a
CC   medicament for treating mammalian bacterial infection. The present
CC   sequence represents a modified antibacterial peptide, which may be used
CC   to produce the composition of the invention
XX
SQ   Sequence 18 AA;

  AAG62740  Length: 18  March 11, 2004 17:24  Type: P  Check: 4080  ..

       1  DKGXXLPRPT PPRPIYXX
```

```
!!AA_SEQUENCE 1.0
ID   AAG62734 standard; peptide; 20 AA.
XX
AC   AAG62734;
XX
DT   17-SEP-2001  (first entry)
XX
DE   Amino acid sequence of antibacterial peptide pyrrhocoricin.
XX
KW   Multi-helical lid; heat shock protein; hsp; protein folding;
KW   pathogenic infection; bacterial infection; antibacterial.
XX
OS   Unidentified.
XX
PN   WO200153509-A2.
XX
PD   26-JUL-2001.
XX
PF   19-JAN-2001; 2001WO-US001812.
XX
PR   21-JAN-2000; 2000US-0177565P.
PR   03-OCT-2000; 2000US-0237599P.
XX
PA   (WIST-) WISTAR INST ANATOMY & BIOLOGY.
PA   (UYCR-) UNIV CREIGHTON.
XX
PI   Otvos L,  Blaszczyk-Thurin M,  Rogers M,  Lovas S;
XX
DR   WPI; 2001-451911/48.
XX
PT   Composition, used to treat a pathogenic infection and eliminate a plant,
PT   insect, or animal pest, comprises a molecule that binds to a heat shock
PT   protein.
XX
```

PS Example 6; Page 64; 124pp; English.
XX
CC The specification describes a composition that comprises a synthetic non-
CC naturally occurring molecule that binds to a selected multi-helical lid
CC of a heat shock protein (hsp) of a selected organism, where the molecule
CC inhibits protein folding activity of the hsp, and a carrier, where
CC exposure of the organism to the composition retards the growth and
CC reproduction of the organism. The composition is used to treat a mammal
CC suffering from a pathogenic infection, in the manufacture of a medicament
CC for treating a mammal for a pathogenic infection, and to eliminate a
CC plant, insect, or animal pest. It is used in the manufacture of a
CC medicament for treating mammalian bacterial infection. The present
CC sequence represents an antibacterial peptide, which may be used to
CC produce the composition of the invention
XX
SQ Sequence 20 AA;

AAG62734 Length: 20 March 11, 2004 17:24 Type: P Check: 6865 ..

1 VDKGSYLPRP TPPRPIYNRN

!!AA_SEQUENCE 1.0
ID AAG62743 standard; peptide; 21 AA.
XX
AC AAG62743;
XX
DT 17-SEP-2001 (first entry)
XX
DE Amino acid sequence of modified antibacterial peptide pyrrhocoricin.
XX
KW Multi-helical lid; heat shock protein; hsp; protein folding;
KW pathogenic infection; bacterial infection; antibacterial.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Modified-site 1
FT /note= "biotin attached"
XX
PN WO200153509-A2.
XX
PD 26-JUL-2001.
XX
PF 19-JAN-2001; 2001WO-US001812.
XX
PR 21-JAN-2000; 2000US-0177565P.
PR 03-OCT-2000; 2000US-0237599P.
XX
PA (WIST-) WISTAR INST ANATOMY & BIOLOGY.
PA (UYCR-) UNIV CREIGHTON.
XX
PI Otvos L, Blaszczyk-Thurin M, Rogers M, Lovas S;
XX
DR WPI; 2001-451911/48.
XX
PT Composition, used to treat a pathogenic infection and eliminate a plant,
PT insect, or animal pest, comprises a molecule that binds to a heat shock
PT protein.
XX
PS Example 1; Page 46; 124pp; English.
XX
CC The specification describes a composition that comprises a synthetic non-
CC naturally occurring molecule that binds to a selected multi-helical lid
CC of a heat shock protein (hsp) of a selected organism, where the molecule
CC inhibits protein folding activity of the hsp, and a carrier, where
CC exposure of the organism to the composition retards the growth and
CC reproduction of the organism. The composition is used to treat a mammal
CC suffering from a pathogenic infection, in the manufacture of a medicament
CC for treating a mammal for a pathogenic infection, and to eliminate a
CC plant, insect, or animal pest. It is used in the manufacture of a
CC medicament for treating mammalian bacterial infection. The present
CC sequence represents a modified antibacterial peptide, which may be used
CC to produce the composition of the invention
XX
SQ Sequence 21 AA;

AAG62743 Length: 21 March 11, 2004 17:24 Type: P Check: 8536 ..

1 KVDKGSYLPR PTPPRPIYNR N

!!AA_SEQUENCE 1.0
ID AAG62756 standard; peptide; 21 AA.
XX
AC AAG62756;
XX
DT 17-SEP-2001 (first entry)
XX
DE Amino acid sequence of modified antibacterial peptide pyrrhocoricin.
XX
KW Multi-helical lid; heat shock protein; hsp; protein folding;
KW pathogenic infection; bacterial infection; antibacterial.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Modified-site 1
FT /note= "fluorescein attached"
XX
PN WO200153509-A2.
XX
PD 26-JUL-2001.
XX
PF 19-JAN-2001; 2001WO-US001812.
XX
PR 21-JAN-2000; 2000US-0177565P.
PR 03-OCT-2000; 2000US-0237599P.
XX
PA (WIST-) WISTAR INST ANATOMY & BIOLOGY.
PA (UYCR-) UNIV CREIGHTON.
XX
PI Otvos L, Blaszczyk-Thurin M, Rogers M, Lovas S;
XX
DR WPI; 2001-451911/48.
XX
PT Composition, used to treat a pathogenic infection and eliminate a plant,
PT insect, or animal pest, comprises a molecule that binds to a heat shock
PT protein.
XX
PS Example 1; Page 56; 124pp; English.
XX
CC The specification describes a composition that comprises a synthetic non-
CC naturally occurring molecule that binds to a selected multi-helical lid
CC of a heat shock protein (hsp) of a selected organism, where the molecule
CC inhibits protein folding activity of the hsp, and a carrier, where
CC exposure of the organism to the composition retards the growth and
CC reproduction of the organism. The composition is used to treat a mammal
CC suffering from a pathogenic infection, in the manufacture of a medicament
CC for treating a mammal for a pathogenic infection, and to eliminate a
CC plant, insect, or animal pest. It is used in the manufacture of a
CC medicament for treating mammalian bacterial infection. The present
CC sequence represents a modified antibacterial peptide, which may be used
CC to produce the composition of the invention
XX
SQ Sequence 21 AA;

AAG62756 Length: 21 March 11, 2004 17:24 Type: P Check: 8536 ..

1 KVDKGSYLPR PTPPRPIYNR N

!!AA_SEQUENCE 1.0
ID AAY72457 standard; peptide; 20 AA.
XX
AC AAY72457;
XX
DT 06-AUG-2003 (revised)
DT 24-APR-2001 (first entry)

```
XX
DE   Pyrrhocoricin-modified Peptide 13.
XX
KW   Pyrrhocoricin-derived peptide; antibacterial; fungicidal; therapy;
KW   fungal infection; bacterial infection; candidiasis; drug development.
XX
OS   Pyrrhocoris apterus.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 1. .20
FT                   /note= "D-form residues"
XX
PN   WO200078956-A1.
XX
PD   28-DEC-2000.
XX
PF   21-JUN-2000; 2000WO-US016989.
XX
PR   23-JUN-1999;   99US-0140606P.
PR   15-SEP-1999;   99US-0154135P.
XX
PA   (WIST-) WISTAR INST ANATOMY & BIOLOGY.
XX
PI   Otvos L;
XX
DR   WPI; 2001-112323/12.
XX
PT   Polypeptides derived from the peptide pyrrhocoricin, useful for treating
PT   fungal infections and Gram negative/positive bacterial infections.
XX
PS   Example 1; Page 25; 75pp; English.
XX
CC   The present peptide sequence is inactive Pyrrhocoricin-modified Peptide
CC   13. Pyrrhocoricin is a glycopeptide characterised by the presence of a
CC   disaccharide in the mid-chain position. The invention relates to
CC   pyrrhocoricin-derived peptides which have anti-bacterial or anti-fungal
CC   activity. These peptides have metabolic stability in mammalian serum. The
CC   pyrrhocoricin-derived peptides are used in the treatment of bacterial
CC   infections caused by Gram positive or Gram negative bacterium and fungal
CC   infections of skin, nails, mucus membranes and intestines e.g.,
CC   candidiasis. These peptides are also useful in anti-bacterial or anti-
CC   fungal pharmaceutical compositions, drug development and identification
CC   of other antibiotic or anti-fungal compounds. (Updated on 06-AUG-2003 to
CC   correct OS field.)
XX
SQ   Sequence 20 AA;

  AAY72457  Length: 20  March 11, 2004 17:24  Type: P  Check: 6865  ..

      1  VDKGSYLPRP TPPRPIYNRN

!!AA_SEQUENCE 1.0
ID   AAY72439 standard; peptide; 21 AA.
XX
AC   AAY72439;
XX
DT   06-AUG-2003  (revised)
DT   24-APR-2001  (first entry)
XX
DE   Pyrrhocoricin-modified Peptide 4.
XX
KW   Pyrrhocoricin-derived peptide; antibacterial; fungicidal; therapy;
KW   fungal infection; bacterial infection; candidiasis; drug development.
XX
OS   Pyrrhocoris apterus.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   1
FT                   /note= "N-terminal acetyl"
XX
PN   WO200078956-A1.
XX
PD   28-DEC-2000.
XX
PF   21-JUN-2000; 2000WO-US016989.
XX
PR   23-JUN-1999;   99US-0140606P.
PR   15-SEP-1999;   99US-0154135P.
XX
PA   (WIST-) WISTAR INST ANATOMY & BIOLOGY.
XX
PI   Otvos L;
XX
DR   WPI; 2001-112323/12.
XX
PT   Polypeptides derived from the peptide pyrrhocoricin, useful for treating
PT   fungal infections and Gram negative/positive bacterial infections.
XX
PS   Claim 23; Page 45; 75pp; English.
XX
CC   The present peptide sequence is active Pyrrhocoricin-modified Peptide 4.
CC   Pyrrhocoricin is a glycopeptide characterised by the presence of a
CC   disaccharide in the mid-chain position. The invention relates to
CC   pyrrhocoricin-derived peptides which have anti-bacterial or anti-fungal
CC   activity. These peptides have metabolic stability in mammalian serum. The
CC   pyrrhocoricin-derived peptides are used in the treatment of bacterial
CC   infections caused by Gram positive or Gram negative bacterium and fungal
CC   infections of skin, nails, mucus membranes and intestines e.g.,
CC   candidiasis. These peptides are also useful in anti-bacterial or anti-
CC   fungal pharmaceutical compositions, drug development and identification
CC   of other antibiotic or anti-fungal compounds. (Updated on 06-AUG-2003 to
CC   correct OS field.)
XX
SQ   Sequence 21 AA;

  AAY72439  Length: 21  March 11, 2004 17:24  Type: P  Check: 8543  ..

      1  RVDKGSYLPR PTPPRPIYNR N

!!AA_SEQUENCE 1.0
ID   AAY72444 standard; peptide; 21 AA.
XX
AC   AAY72444;
XX
DT   06-AUG-2003  (revised)
DT   24-APR-2001  (first entry)
XX
DE   Pyrrhocoricin-modified Peptide 9.
XX
KW   Pyrrhocoricin-derived peptide; antibacterial; fungicidal; therapy;
KW   fungal infection; bacterial infection; candidiasis; drug development.
XX
OS   Pyrrhocoris apterus.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   1
FT                   /note= "N-terminal acetyl"
FT   Modified-site   21
FT                   /note= "Beta-acetyl-2,3-diamino propionic acid"
XX
PN   WO200078956-A1.
XX
PD   28-DEC-2000.
XX
PF   21-JUN-2000; 2000WO-US016989.
XX
PR   23-JUN-1999;   99US-0140606P.
PR   15-SEP-1999;   99US-0154135P.
XX
PA   (WIST-) WISTAR INST ANATOMY & BIOLOGY.
XX
PI   Otvos L;
XX
```

```
DR   WPI; 2001-112323/12.
XX
PT   Polypeptides derived from the peptide pyrrhocoricin, useful for treating
PT   fungal infections and Gram negative/positive bacterial infections.
XX
PS   Claim 27; Page 46; 75pp; English.
XX
CC   The present peptide sequence is active Pyrrhocoricin-modified Peptide 9.
CC   Pyrrhocoricin is a glycopeptide characterised by the presence of a
CC   disaccharide in the mid-chain position. The invention relates to
CC   pyrrhocoricin-derived peptides which have anti-bacterial or anti-fungal
CC   activity. These peptides have metabolic stability in mammalian serum. The
CC   pyrrhocoricin-derived peptides are used in the treatment of bacterial
CC   infections caused by Gram positive or Gram negative bacterium and fungal
CC   infections of skin, nails, mucus membranes and intestines e.g.,
CC   candidiasis. These peptides are also useful in anti-bacterial or anti-
CC   fungal pharmaceutical compositions, drug development and identification
CC   of other antibiotic or anti-fungal compounds. (Updated on 06-AUG-2003 to
CC   correct OS field.)
XX
SQ   Sequence 21 AA;

  AAY72444  Length: 21  March 11, 2004 17:24  Type: P  Check: 8746  ..

       1  KVDKGSYLPR PTPPRPIYNR X

!!AA_SEQUENCE 1.0
ID   AAY72454 standard; peptide; 21 AA.
XX
AC   AAY72454;
XX
DT   06-AUG-2003  (revised)
DT   24-APR-2001  (first entry)
XX
DE   Pyrrhocoricin-modified Peptide 22.
XX
KW   Pyrrhocoricin-derived peptide; antibacterial; fungicidal; therapy;
KW   fungal infection; bacterial infection; candidiasis; drug development.
XX
OS   Pyrrhocoris apterus.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   1
FT                   /note= "N-terminal acetyl"
FT   Modified-site   21
FT                   /note= "Beta-acetyl-2,3-diamino propionic acid"
XX
PN   WO200078956-A1.
XX
PD   28-DEC-2000.
XX
PF   21-JUN-2000; 2000WO-US016989.
XX
PR   23-JUN-1999;   99US-0140606P.
PR   15-SEP-1999;   99US-0154135P.
XX
PA   (WIST-) WISTAR INST ANATOMY & BIOLOGY.
XX
PI   Otvos L;
XX
DR   WPI; 2001-112323/12.
XX
PT   Polypeptides derived from the peptide pyrrhocoricin, useful for treating
PT   fungal infections and Gram negative/positive bacterial infections.
XX
PS   Claim 35; Page 47; 75pp; English.
XX
CC   The present peptide sequence is active Pyrrhocoricin-modified Peptide 22.
CC   Pyrrhocoricin is a glycopeptide characterised by the presence of a
CC   disaccharide in the mid-chain position. The invention relates to
CC   pyrrhocoricin-derived peptides which have anti-bacterial or anti-fungal
CC   activity. These peptides have metabolic stability in mammalian serum. The
CC   pyrrhocoricin-derived peptides are used in the treatment of bacterial
CC   infections caused by Gram positive or Gram negative bacterium and fungal
CC   infections of skin, nails, mucus membranes and intestines e.g.,
CC   candidiasis. These peptides are also useful in anti-bacterial or anti-
CC   fungal pharmaceutical compositions, drug development and identification
CC   of other antibiotic or anti-fungal compounds. (Updated on 06-AUG-2003 to
CC   correct OS field.)
XX
SQ   Sequence 21 AA;

  AAY72454  Length: 21  March 11, 2004 17:24  Type: P  Check: 8753  ..

       1  RVDKGSYLPR PTPPRPIYNR X

!!AA_SEQUENCE 1.0
ID   AAY72455 standard; peptide; 20 AA.
XX
AC   AAY72455;
XX
DT   06-AUG-2003  (revised)
DT   24-APR-2001  (first entry)
XX
DE   Pyrrhocoricin-modified Peptide 23.
XX
KW   Pyrrhocoricin-derived peptide; antibacterial; fungicidal; therapy;
KW   fungal infection; bacterial infection; candidiasis; drug development.
XX
OS   Pyrrhocoris apterus.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   1
FT                   /note= "Homoproline or 1-aminocyclo-hexane carboxylic
FT                   acid"
FT   Misc-difference 5
FT                   /note= "Wild type Ser substituted with Ala"
FT   Misc-difference 6
FT                   /note= "Wild type Tyr substituted with Phe"
FT   Modified-site   20
FT                   /note= "Beta-acetyl-2,3-diamino propionic acid"
XX
PN   WO200078956-A1.
XX
PD   28-DEC-2000.
XX
PF   21-JUN-2000; 2000WO-US016989.
XX
PR   23-JUN-1999;   99US-0140606P.
PR   15-SEP-1999;   99US-0154135P.
XX
PA   (WIST-) WISTAR INST ANATOMY & BIOLOGY.
XX
PI   Otvos L;
XX
DR   WPI; 2001-112323/12.
XX
PT   Polypeptides derived from the peptide pyrrhocoricin, useful for treating
PT   fungal infections and Gram negative/positive bacterial infections.
XX
PS   Example 1; Page 28; 75pp; English.
XX
CC   The present peptide sequence is inactive Pyrrhocoricin-modified Peptide
CC   23. Pyrrhocoricin is a glycopeptide characterised by the presence of a
CC   disaccharide in the mid-chain position. The invention relates to
CC   pyrrhocoricin-derived peptides which have anti-bacterial or anti-fungal
CC   activity. These peptides have metabolic stability in mammalian serum. The
CC   pyrrhocoricin-derived peptides are used in the treatment of bacterial
CC   infections caused by Gram positive or Gram negative bacterium and fungal
CC   infections of skin, nails, mucus membranes and intestines e.g.,
CC   candidiasis. These peptides are also useful in anti-bacterial or anti-
CC   fungal pharmaceutical compositions, drug development and identification
CC   of other antibiotic or anti-fungal compounds. (Updated on 06-AUG-2003 to
CC   correct OS field.)
```

```
XX
SQ   Sequence 20 AA;

  AAY72455  Length: 20  March 11, 2004 17:24  Type: P  Check: 6863  ..

       1  XDKGAFLPRP TPPRPIYNRX

!!AA_SEQUENCE 1.0
ID   AAY72461 standard; peptide; 23 AA.
XX
AC   AAY72461;
XX
DT   06-AUG-2003  (revised)
DT   24-APR-2001  (first entry)
XX
DE   Pyrrhocoricin-modified peptide.
XX
KW   Pyrrhocoricin-derived peptide; antibacterial; fungicidal; therapy;
KW   fungal infection; bacterial infection; candidiasis; drug development.
XX
OS   Pyrrhocoris apterus.
OS   Synthetic.
XX
PN   WO200078956-A1.
XX
PD   28-DEC-2000.
XX
PF   21-JUN-2000; 2000WO-US016989.
XX
PR   23-JUN-1999;   99US-0140606P.
PR   15-SEP-1999;   99US-0154135P.
XX
PA   (WIST-) WISTAR INST ANATOMY & BIOLOGY.
XX
PI   Otvos L;
XX
DR   WPI; 2001-112323/12.
XX
PT   Polypeptides derived from the peptide pyrrhocoricin, useful for treating
PT   fungal infections and Gram negative/positive bacterial infections.
XX
PS   Claim 21; Page 45; 75pp; English.
XX
CC   The present peptide sequence is inactive Pyrrhocoricin-modified peptide.
CC   Pyrrhocoricin is a glycopeptide characterised by the presence of a
CC   disaccharide in the mid-chain position. The invention relates to
CC   pyrrhocoricin-derived peptides which have anti-bacterial or anti-fungal
CC   activity. These peptides have metabolic stability in mammalian serum. The
CC   pyrrhocoricin-derived peptides are used in the treatment of bacterial
CC   infections caused by Gram positive or Gram negative bacterium and fungal
CC   infections of skin, nails, mucus membranes and intestines e.g.,
CC   candidiasis. These peptides are also useful in anti-bacterial or anti-
CC   fungal pharmaceutical compositions, drug development and identification
CC   of other antibiotic or anti-fungal compounds. (Updated on 06-AUG-2003 to
CC   correct OS field.)
XX
SQ   Sequence 23 AA;

  AAY72461  Length: 23  March 11, 2004 17:24  Type: P  Check: 2100  ..

       1  VDKVDKGSYL PRPTPPRPIY NRN

!!AA_SEQUENCE 1.0
ID   AAY72442 standard; peptide; 20 AA.
XX
AC   AAY72442;
XX
DT   06-AUG-2003  (revised)
DT   24-APR-2001  (first entry)
XX
DE   Pyrrhocoricin-modified Peptide 7.
XX
KW   Pyrrhocoricin-derived peptide; antibacterial; fungicidal; therapy;
KW   fungal infection; bacterial infection; candidiasis; drug development.
XX
OS   Pyrrhocoris apterus.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   1
FT                   /note= "N-terminal acetyl"
FT   Modified-site   11
FT                   /note= "Modified with galactose-2-acetamido-2-deoxy-
FT                   galactose (Gal-GalNAc)"
XX
PN   WO200078956-A1.
XX
PD   28-DEC-2000.
XX
PF   21-JUN-2000; 2000WO-US016989.
XX
PR   23-JUN-1999;   99US-0140606P.
PR   15-SEP-1999;   99US-0154135P.
XX
PA   (WIST-) WISTAR INST ANATOMY & BIOLOGY.
XX
PI   Otvos L;
XX
DR   WPI; 2001-112323/12.
XX
PT   Polypeptides derived from the peptide pyrrhocoricin, useful for treating
PT   fungal infections and Gram negative/positive bacterial infections.
XX
PS   Example 1; Page 24; 75pp; English.
XX
CC   The present peptide sequence is inactive Pyrrhocoricin-modified Peptide
CC   7. Pyrrhocoricin is a glycopeptide characterised by the presence of a
CC   disaccharide in the mid-chain position. The invention relates to
CC   pyrrhocoricin-derived peptides which have anti-bacterial or anti-fungal
CC   activity. These peptides have metabolic stability in mammalian serum. The
CC   pyrrhocoricin-derived peptides are used in the treatment of bacterial
CC   infections caused by Gram positive or Gram negative bacterium and fungal
CC   infections of skin, nails, mucus membranes and intestines e.g.,
CC   candidiasis. These peptides are also useful in anti-bacterial or anti-
CC   fungal pharmaceutical compositions, drug development and identification
CC   of other antibiotic or anti-fungal compounds. (Updated on 06-AUG-2003 to
CC   correct OS field.)
XX
SQ   Sequence 20 AA;

  AAY72442  Length: 20  March 11, 2004 17:24  Type: P  Check: 6865  ..

       1  VDKGSYLPRP TPPRPIYNRN

!!AA_SEQUENCE 1.0
ID   AAY72448 standard; peptide; 21 AA.
XX
AC   AAY72448;
XX
DT   06-AUG-2003  (revised)
DT   24-APR-2001  (first entry)
XX
DE   Pyrrhocoricin-modified Peptide 16.
XX
KW   Pyrrhocoricin-derived peptide; antibacterial; fungicidal; therapy;
KW   fungal infection; bacterial infection; candidiasis; drug development;
KW   cyclic.
XX
OS   Pyrrhocoris apterus.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   1
FT                   /note= "Forms a cyclic linkage with Asp at the C-terminal
FT                   end"
FT   Modified-site   21
```

```
FT                  /note= "Forms a cyclic linkage with Lys at the N-terminal
FT                  end"
XX
PN   WO200078956-A1.
XX
PD   28-DEC-2000.
XX
PF   21-JUN-2000; 2000WO-US016989.
XX
PR   23-JUN-1999;   99US-0140606P.
PR   15-SEP-1999;   99US-0154135P.
XX
PA   (WIST-) WISTAR INST ANATOMY & BIOLOGY.
XX
PI   Otvos L;
XX
DR   WPI; 2001-112323/12.
XX
PT   Polypeptides derived from the peptide pyrrhocoricin, useful for treating
PT   fungal infections and Gram negative/positive bacterial infections.
XX
PS   Example 1; Page 26; 75pp; English.
XX
CC   The present peptide sequence is inactive Pyrrhocoricin-modified Peptide
CC   16. Pyrrhocoricin is a glycopeptide characterised by the presence of a
CC   disaccharide in the mid-chain position. The invention relates to
CC   pyrrhocoricin-derived peptides which have anti-bacterial or anti-fungal
CC   activity. These peptides have metabolic stability in mammalian serum. The
CC   pyrrhocoricin-derived peptides are used in the treatment of bacterial
CC   infections caused by Gram positive or Gram negative bacterium and fungal
CC   infections of skin, nails, mucus membranes and intestines e.g.,
CC   candidiasis. These peptides are also useful in anti-bacterial or anti-
CC   fungal pharmaceutical compositions, drug development and identification
CC   of other antibiotic or anti-fungal compounds. (Updated on 06-AUG-2003 to
CC   correct OS field.)
XX
SQ   Sequence 21 AA;

  AAY72448  Length: 21  March 11, 2004 17:24  Type: P  Check: 8326  ..

       1  KVDKGSYLPR PTPPRPIYNR D

!!AA_SEQUENCE 1.0
ID   AAY72449 standard; peptide; 29 AA.
XX
AC   AAY72449;
XX
DT   06-AUG-2003  (revised)
DT   24-APR-2001  (first entry)
XX
DE   Pyrrhocoricin-modified Peptide 17.
XX
KW   Pyrrhocoricin-derived peptide; antibacterial; fungicidal; therapy;
KW   fungal infection; bacterial infection; candidiasis; drug development;
KW   cyclic.
XX
OS   Pyrrhocoris apterus.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   1
FT                   /note= "Forms a cyclic linkage with Asn at the C-terminal
FT                   end"
FT   Modified-site   29
FT                   /note= "Forms a cyclic linkage with Arg at the N-terminal
FT                   end"
XX
PN   WO200078956-A1.
XX
PD   28-DEC-2000.
XX
PF   21-JUN-2000; 2000WO-US016989.
XX
PR   23-JUN-1999;   99US-0140606P.
PR   15-SEP-1999;   99US-0154135P.
XX
PA   (WIST-) WISTAR INST ANATOMY & BIOLOGY.
XX
PI   Otvos L;
XX
DR   WPI; 2001-112323/12.
XX
PT   Polypeptides derived from the peptide pyrrhocoricin, useful for treating
PT   fungal infections and Gram negative/positive bacterial infections.
XX
PS   Claim 37; Page 47; 75pp; English.
XX
CC   The present peptide sequence is active Pyrrhocoricin-modified Peptide 17.
CC   This cyclic non-glycosylated peptide is the most active peptide.
CC   Pyrrhocoricin is a glycopeptide characterised by the presence of a
CC   disaccharide in the mid-chain position. The invention relates to
CC   pyrrhocoricin-derived peptides which have anti-bacterial or anti-fungal
CC   activity. These peptides have metabolic stability in mammalian serum. The
CC   pyrrhocoricin-derived peptides are used in the treatment of bacterial
CC   infections caused by Gram positive or Gram negative bacterium and fungal
CC   infections of skin, nails, mucus membranes and intestines e.g.,
CC   candidiasis. These peptides are also useful in anti-bacterial or anti-
CC   fungal pharmaceutical compositions, drug development and identification
CC   of other antibiotic or anti-fungal compounds. (Updated on 06-AUG-2003 to
CC   correct OS field.)
XX
SQ   Sequence 29 AA;

  AAY72449  Length: 29  March 11, 2004 17:24  Type: P  Check: 4782  ..

       1  RPPTPRPLKV DKGSYLPRPT PPRPIYNRN

!!AA_SEQUENCE 1.0
ID   AAY72424 standard; peptide; 18 AA.
XX
AC   AAY72424;
XX
DT   06-AUG-2003  (revised)
DT   24-APR-2001  (first entry)
XX
DE   Pyrrhocoricin based generic peptide #1.
XX
KW   Pyrrhocoricin-derived peptide; antibacterial; fungicidal; therapy;
KW   fungal infection; bacterial infection; candidiasis; drug development.
XX
OS   Pyrrhocoris apterus.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   1
FT                   /note= "Optionally attached to additional amino acids or
FT                   modified with a straight chain, branched, cyclic or
FT                   heterocyclic alkyl group (preferably 1-aminocyclo-hexane
FT                   carboxylic acid), heterocyclic alkanoyl group or a
FT                   positively charged reporter group (preferably biotin,
FT                   5(6) carboxyfluorescein)"
FT   Misc-difference 4
FT                   /note= "Ser or any amino acid"
FT   Misc-difference 5
FT                   /note= "Tyr or any amino acid"
FT   Misc-difference 17
FT                   /note= "Asn or any amino acid"
FT   Modified-site   18
FT                   /note= "Optionally attached to additional amino acids or
FT                   modified with an amide, an imide or a sugar moiety"
FT   Misc-difference 18
FT                   /note= "Arg or any amino acid"
XX
PN   WO200078956-A1.
XX
PD   28-DEC-2000.
```

XX
PF 21-JUN-2000; 2000WO-US016989.
XX
PR 23-JUN-1999; 99US-0140606P.
PR 15-SEP-1999; 99US-0154135P.
XX
PA (WIST-) WISTAR INST ANATOMY & BIOLOGY.
XX
PI Otvos L;
XX
DR WPI; 2001-112323/12.
XX
PT Polypeptides derived from the peptide pyrrhocoricin, useful for treating
PT fungal infections and Gram negative/positive bacterial infections.
XX
PS Claim 1; Page 42; 75pp; English.
XX
CC The present sequence is a pyrrhocoricin based generic peptide which has
CC anti-bacterial or anti-fungal activity. Pyrrhocoricin is a glycopeptide
CC characterised by the presence of a disaccharide in the mid-chain
CC position. The invention relates to pyrrhocoricin-derived peptides. These
CC peptides have metabolic stability in mammalian serum. The pyrrhocoricin-
CC derived peptides are used in the treatment of bacterial infections caused
CC by Gram positive or Gram negative bacterium and fungal infections of
CC skin, nails, mucus membranes and intestines e.g., candidiasis. These
CC peptides are also useful in anti-bacterial or anti-fungal pharmaceutical
CC compositions, drug development and identification of other antibiotic or
CC anti-fungal compounds. (Updated on 06-AUG-2003 to correct OS field.)
XX
SQ Sequence 18 AA;

AAY72424 Length: 18 March 11, 2004 17:24 Type: P Check: 4080 ..

1 DKGXXLPRPT PPRPIYXX

!!AA_SEQUENCE 1.0
ID AAY72440 standard; peptide; 21 AA.
XX
AC AAY72440;
XX
DT 06-AUG-2003 (revised)
DT 24-APR-2001 (first entry)
XX
DE Pyrrhocoricin-modified Peptide 5.
XX
KW Pyrrhocoricin-derived peptide; antibacterial; fungicidal; therapy;
KW fungal infection; bacterial infection; candidiasis; drug development.
XX
OS Pyrrhocoris apterus.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Modified-site 1
FT /note= "N-terminal acetyl"
XX
PN WO200078956-A1.
XX
PD 28-DEC-2000.
XX
PF 21-JUN-2000; 2000WO-US016989.
XX
PR 23-JUN-1999; 99US-0140606P.
PR 15-SEP-1999; 99US-0154135P.
XX
PA (WIST-) WISTAR INST ANATOMY & BIOLOGY.
XX
PI Otvos L;
XX
DR WPI; 2001-112323/12.
XX
PT Polypeptides derived from the peptide pyrrhocoricin, useful for treating
PT fungal infections and Gram negative/positive bacterial infections.
XX
PS Claim 24; Page 45; 75pp; English.
XX
CC The present peptide sequence is active Pyrrhocoricin-modified Peptide 5.
CC Pyrrhocoricin is a glycopeptide characterised by the presence of a
CC disaccharide in the mid-chain position. The invention relates to
CC pyrrhocoricin-derived peptides which have anti-bacterial or anti-fungal
CC activity. These peptides have metabolic stability in mammalian serum. The
CC pyrrhocoricin-derived peptides are used in the treatment of bacterial
CC infections caused by Gram positive or Gram negative bacterium and fungal
CC infections of skin, nails, mucus membranes and intestines e.g.,
CC candidiasis. These peptides are also useful in anti-bacterial or anti-
CC fungal pharmaceutical compositions, drug development and identification
CC of other antibiotic or anti-fungal compounds. (Updated on 06-AUG-2003 to
CC correct OS field.)
XX
SQ Sequence 21 AA;

AAY72440 Length: 21 March 11, 2004 17:24 Type: P Check: 8536 ..

1 KVDKGSYLPR PTPPRPIYNR N

!!AA_SEQUENCE 1.0
ID AAY72441 standard; peptide; 20 AA.
XX
AC AAY72441;
XX
DT 06-AUG-2003 (revised)
DT 24-APR-2001 (first entry)
XX
DE Pyrrhocoricin-modified Peptide 6.
XX
KW Pyrrhocoricin-derived peptide; antibacterial; fungicidal; therapy;
KW fungal infection; bacterial infection; candidiasis; drug development.
XX
OS Pyrrhocoris apterus.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Modified-site 1
FT /note= "Homoproline or 1-aminocyclo-hexane carboxylic
FT acid"
XX
PN WO200078956-A1.
XX
PD 28-DEC-2000.
XX
PF 21-JUN-2000; 2000WO-US016989.
XX
PR 23-JUN-1999; 99US-0140606P.
PR 15-SEP-1999; 99US-0154135P.
XX
PA (WIST-) WISTAR INST ANATOMY & BIOLOGY.
XX
PI Otvos L;
XX
DR WPI; 2001-112323/12.
XX
PT Polypeptides derived from the peptide pyrrhocoricin, useful for treating
PT fungal infections and Gram negative/positive bacterial infections.
XX
PS Claim 25; Page 45; 75pp; English.
XX
CC The present peptide sequence is active Pyrrhocoricin-modified Peptide 6.
CC Pyrrhocoricin is a glycopeptide characterised by the presence of a
CC disaccharide in the mid-chain position. The invention relates to
CC pyrrhocoricin-derived peptides which have anti-bacterial or anti-fungal
CC activity. These peptides have metabolic stability in mammalian serum. The
CC pyrrhocoricin-derived peptides are used in the treatment of bacterial
CC infections caused by Gram positive or Gram negative bacterium and fungal
CC infections of skin, nails, mucus membranes and intestines e.g.,
CC candidiasis. These peptides are also useful in anti-bacterial or anti-
CC fungal pharmaceutical compositions, drug development and identification
CC of other antibiotic or anti-fungal compounds. (Updated on 06-AUG-2003 to

```
CC   correct OS field.)
XX
SQ   Sequence 20 AA;

  AAY72441  Length: 20  March 11, 2004 17:24  Type: P  Check: 6867  ..

       1  XDKGSYLPRP TPPRPIYNRN

!!AA_SEQUENCE 1.0
ID   AAY72443 standard; peptide; 20 AA.
XX
AC   AAY72443;
XX
DT   06-AUG-2003  (revised)
DT   24-APR-2001  (first entry)
XX
DE   Pyrrhocoricin-modified Peptide 8.
XX
KW   Pyrrhocoricin-derived peptide; antibacterial; fungicidal; therapy;
KW   fungal infection; bacterial infection; candidiasis; drug development.
XX
OS   Pyrrhocoris apterus.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   1
FT                   /note= "N-terminal acetyl"
FT   Modified-site   20
FT                   /note= "C-terminal imide"
XX
PN   WO200078956-A1.
XX
PD   28-DEC-2000.
XX
PF   21-JUN-2000; 2000WO-US016989.
XX
PR   23-JUN-1999;   99US-0140606P.
PR   15-SEP-1999;   99US-0154135P.
XX
PA   (WIST-) WISTAR INST ANATOMY & BIOLOGY.
XX
PI   Otvos L;
XX
DR   WPI; 2001-112323/12.
XX
PT   Polypeptides derived from the peptide pyrrhocoricin, useful for treating
PT   fungal infections and Gram negative/positive bacterial infections.
XX
PS   Claim 26; Page 45; 75pp; English.
XX
CC   The present peptide sequence is active Pyrrhocoricin-modified Peptide 8.
CC   Pyrrhocoricin is a glycopeptide characterised by the presence of a
CC   disaccharide in the mid-chain position. The invention relates to
CC   pyrrhocoricin-derived peptides which have anti-bacterial or anti-fungal
CC   activity. These peptides have metabolic stability in mammalian serum. The
CC   pyrrhocoricin-derived peptides are used in the treatment of bacterial
CC   infections caused by Gram positive or Gram negative bacterium and fungal
CC   infections of skin, nails, mucus membranes and intestines e.g.,
CC   candidiasis. These peptides are also useful in anti-bacterial or anti-
CC   fungal pharmaceutical compositions, drug development and identification
CC   of other antibiotic or anti-fungal compounds. (Updated on 06-AUG-2003 to
CC   correct OS field.)
XX
SQ   Sequence 20 AA;

  AAY72443  Length: 20  March 11, 2004 17:24  Type: P  Check: 6898  ..

       1  KVDKGSYLPR PTPPRPIYNR

!!AA_SEQUENCE 1.0
ID   AAY72447 standard; peptide; 20 AA.
XX
AC   AAY72447;
XX
DT   06-AUG-2003  (revised)
DT   24-APR-2001  (first entry)
XX
DE   Pyrrhocoricin-modified Peptide 12.
XX
KW   Pyrrhocoricin-derived peptide; antibacterial; fungicidal; therapy;
KW   fungal infection; bacterial infection; candidiasis; drug development.
XX
OS   Pyrrhocoris apterus.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 1
FT                   /note= "D-form residue"
FT   Misc-difference 20
FT                   /note= "D-form residue"
XX
PN   WO200078956-A1.
XX
PD   28-DEC-2000.
XX
PF   21-JUN-2000; 2000WO-US016989.
XX
PR   23-JUN-1999;   99US-0140606P.
PR   15-SEP-1999;   99US-0154135P.
XX
PA   (WIST-) WISTAR INST ANATOMY & BIOLOGY.
XX
PI   Otvos L;
XX
DR   WPI; 2001-112323/12.
XX
PT   Polypeptides derived from the peptide pyrrhocoricin, useful for treating
PT   fungal infections and Gram negative/positive bacterial infections.
XX
PS   Claim 30; Page 46; 75pp; English.
XX
CC   The present peptide sequence is active Pyrrhocoricin-modified Peptide 12.
CC   Pyrrhocoricin is a glycopeptide characterised by the presence of a
CC   disaccharide in the mid-chain position. The invention relates to
CC   pyrrhocoricin-derived peptides which have anti-bacterial or anti-fungal
CC   activity. These peptides have metabolic stability in mammalian serum. The
CC   pyrrhocoricin-derived peptides are used in the treatment of bacterial
CC   infections caused by Gram positive or Gram negative bacterium and fungal
CC   infections of skin, nails, mucus membranes and intestines e.g.,
CC   candidiasis. These peptides are also useful in anti-bacterial or anti-
CC   fungal pharmaceutical compositions, drug development and identification
CC   of other antibiotic or anti-fungal compounds. (Updated on 06-AUG-2003 to
CC   correct OS field.)
XX
SQ   Sequence 20 AA;

  AAY72447  Length: 20  March 11, 2004 17:24  Type: P  Check: 6865  ..

       1  VDKGSYLPRP TPPRPIYNRN

!!AA_SEQUENCE 1.0
ID   AAY72453 standard; peptide; 20 AA.
XX
AC   AAY72453;
XX
DT   06-AUG-2003  (revised)
DT   24-APR-2001  (first entry)
XX
DE   Pyrrhocoricin-modified Peptide 21.
XX
KW   Pyrrhocoricin-derived peptide; antibacterial; fungicidal; therapy;
KW   fungal infection; bacterial infection; candidiasis; drug development.
XX
OS   Pyrrhocoris apterus.
OS   Synthetic.
XX
```

```
FH   Key             Location/Qualifiers
FT   Modified-site   1
FT                   /note= "Homoproline or 1-aminocyclo-hexane carboxylic
FT                   acid"
FT   Modified-site   20
FT                   /note= "Beta-acetyl-2,3-diamino propionic acid"
XX
PN   WO200078956-A1.
XX
PD   28-DEC-2000.
XX
PF   21-JUN-2000; 2000WO-US016989.
XX
PR   23-JUN-1999;   99US-0140606P.
PR   15-SEP-1999;   99US-0154135P.
XX
PA   (WIST-) WISTAR INST ANATOMY & BIOLOGY.
XX
PI   Otvos L;
XX
DR   WPI; 2001-112323/12.
XX
PT   Polypeptides derived from the peptide pyrrhocoricin, useful for treating
PT   fungal infections and Gram negative/positive bacterial infections.
XX
PS   Claim 34; Page 47; 75pp; English.
XX
CC   The present peptide sequence is active Pyrrhocoricin-modified Peptide 21.
CC   Pyrrhocoricin is a glycopeptide characterised by the presence of a
CC   disaccharide in the mid-chain position. The invention relates to
CC   pyrrhocoricin-derived peptides which have anti-bacterial or anti-fungal
CC   activity. These peptides have metabolic stability in mammalian serum. The
CC   pyrrhocoricin-derived peptides are used in the treatment of bacterial
CC   infections caused by Gram positive or Gram negative bacterium and fungal
CC   infections of skin, nails, mucus membranes and intestines e.g.,
CC   candidiasis. These peptides are also useful in anti-bacterial or anti-
CC   fungal pharmaceutical compositions, drug development and identification
CC   of other antibiotic or anti-fungal compounds. (Updated on 06-AUG-2003 to
CC   correct OS field.)
XX
SQ   Sequence 20 AA;

  AAY72453  Length: 20  March 11, 2004 17:24  Type: P  Check: 7067  ..

       1  XDKGSYLPRP TPPRPIYNRX

!!AA_SEQUENCE 1.0
ID   AAY72451 standard; peptide; 21 AA.
XX
AC   AAY72451;
XX
DT   06-AUG-2003  (revised)
DT   24-APR-2001  (first entry)
XX
DE   Pyrrhocoricin-modified Peptide 19.
XX
KW   Pyrrhocoricin-derived peptide; antibacterial; fungicidal; therapy;
KW   fungal infection; bacterial infection; candidiasis; drug development.
XX
OS   Pyrrhocoris apterus.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   1
FT                   /note= "N-terminal 5(6)-carboxyfluorescein"
XX
PN   WO200078956-A1.
XX
PD   28-DEC-2000.
XX
PF   21-JUN-2000; 2000WO-US016989.
XX
PR   23-JUN-1999;   99US-0140606P.
PR   15-SEP-1999;   99US-0154135P.
XX
PA   (WIST-) WISTAR INST ANATOMY & BIOLOGY.
XX
PI   Otvos L;
XX
DR   WPI; 2001-112323/12.
XX
PT   Polypeptides derived from the peptide pyrrhocoricin, useful for treating
PT   fungal infections and Gram negative/positive bacterial infections.
XX
PS   Claim 32; Page 46; 75pp; English.
XX
CC   The present peptide sequence is active Pyrrhocoricin-modified Peptide 19.
CC   Pyrrhocoricin is a glycopeptide characterised by the presence of a
CC   disaccharide in the mid-chain position. The invention relates to
CC   pyrrhocoricin-derived peptides which have anti-bacterial or anti-fungal
CC   activity. These peptides have metabolic stability in mammalian serum. The
CC   pyrrhocoricin-derived peptides are used in the treatment of bacterial
CC   infections caused by Gram positive or Gram negative bacterium and fungal
CC   infections of skin, nails, mucus membranes and intestines e.g.,
CC   candidiasis. These peptides are also useful in anti-bacterial or anti-
CC   fungal pharmaceutical compositions, drug development and identification
CC   of other antibiotic or anti-fungal compounds. (Updated on 06-AUG-2003 to
CC   correct OS field.)
XX
SQ   Sequence 21 AA;

  AAY72451  Length: 21  March 11, 2004 17:24  Type: P  Check: 8536  ..

       1  KVDKGSYLPR PTPPRPIYNR N

!!AA_SEQUENCE 1.0
ID   AAY72452 standard; peptide; 21 AA.
XX
AC   AAY72452;
XX
DT   06-AUG-2003  (revised)
DT   24-APR-2001  (first entry)
XX
DE   Pyrrhocoricin-modified Peptide 20.
XX
KW   Pyrrhocoricin-derived peptide; antibacterial; fungicidal; therapy;
KW   fungal infection; bacterial infection; candidiasis; drug development.
XX
OS   Pyrrhocoris apterus.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   1
FT                   /note= "N-terminal acetyl"
FT   Misc-difference 21
FT                   /note= "Wild type Asn substituted with Asp"
XX
PN   WO200078956-A1.
XX
PD   28-DEC-2000.
XX
PF   21-JUN-2000; 2000WO-US016989.
XX
PR   23-JUN-1999;   99US-0140606P.
PR   15-SEP-1999;   99US-0154135P.
XX
PA   (WIST-) WISTAR INST ANATOMY & BIOLOGY.
XX
PI   Otvos L;
XX
DR   WPI; 2001-112323/12.
XX
PT   Polypeptides derived from the peptide pyrrhocoricin, useful for treating
PT   fungal infections and Gram negative/positive bacterial infections.
XX
PS   Claim 33; Page 46; 75pp; English.
```

```
XX
CC   The present peptide sequence is weakly active Pyrrhocoricin-modified
CC   Peptide 20. Pyrrhocoricin is a glycopeptide characterised by the presence
CC   of a disaccharide in the mid-chain position. The invention relates to
CC   pyrrhocoricin-derived peptides which have anti-bacterial or anti-fungal
CC   activity. These peptides have metabolic stability in mammalian serum. The
CC   pyrrhocoricin-derived peptides are used in the treatment of bacterial
CC   infections caused by Gram positive or Gram negative bacterium and fungal
CC   infections of skin, nails, mucus membranes and intestines e.g.,
CC   candidiasis. These peptides are also useful in anti-bacterial or anti-
CC   fungal pharmaceutical compositions, drug development and identification
CC   of other antibiotic or anti-fungal compounds. (Updated on 06-AUG-2003 to
CC   correct OS field.)
XX
SQ   Sequence 21 AA;

  AAY72452  Length: 21  March 11, 2004 17:24  Type: P  Check: 8326  ..

       1  KVDKGSYLPR PTPPRPIYNR D

!!AA_SEQUENCE 1.0
ID   AAY72498 standard; peptide; 20 AA.
XX
AC   AAY72498;
XX
DT   06-AUG-2003  (revised)
DT   24-APR-2001  (first entry)
XX
DE   Pyrrhocoricin-modified peptide #2 for multi-peptide construction.
XX
KW   Pyrrhocoricin-derived peptide; antibacterial; fungicidal; therapy;
KW   fungal infection; bacterial infection; candidiasis; drug development.
XX
OS   Pyrrhocoris apterus.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   1
FT                   /note= "Homoproline or 1-aminocyclo-hexane carboxylic
FT                   acid"
FT   Cross-links     20
FT                   /note= "The carboxy group of the 2-amino-3-acetylamino-
FT                   propionic acid residue 20 of AAY72498 is condensed onto
FT                   the side chain amino group of 2,3-diamino propionic acid
FT                   residue 20 of AAY72435 to cross link the two peptides
FT                   into a multipeptide"
FT   Modified-site   20
FT                   /note= "2-amino-3-acetylamino-propionic acid residue"
XX
PN   WO200078956-A1.
XX
PD   28-DEC-2000.
XX
PF   21-JUN-2000; 2000WO-US016989.
XX
PR   23-JUN-1999;   99US-0140606P.
PR   15-SEP-1999;   99US-0154135P.
XX
PA   (WIST-) WISTAR INST ANATOMY & BIOLOGY.
XX
PI   Otvos L;
XX
DR   WPI; 2001-112323/12.
XX
PT   Polypeptides derived from the peptide pyrrhocoricin, useful for treating
PT   fungal infections and Gram negative/positive bacterial infections.
XX
PS   Claim 51; Page 50; 75pp; English.
XX
CC   The present peptide sequence is Pyrrhocoricin-modified peptide used for
CC   multipeptide construction. Pyrrhocoricin is a glycopeptide characterised
CC   by the presence of a disaccharide in the mid-chain position. The
CC   invention relates to pyrrhocoricin-derived peptides which have anti-
CC   bacterial or anti-fungal activity. These peptides have metabolic
CC   stability in mammalian serum. The pyrrhocoricin-derived peptides are used
CC   in the treatment of bacterial infections caused by Gram positive or Gram
CC   negative bacterium and fungal infections of skin, nails, mucus membranes
CC   and intestines e.g., candidiasis. These peptides are also useful in anti-
CC   bacterial or anti-fungal pharmaceutical compositions, drug development
CC   and identification of other antibiotic or anti-fungal compounds. (Updated
CC   on 06-AUG-2003 to correct OS field.)
XX
SQ   Sequence 20 AA;

  AAY72498  Length: 20  March 11, 2004 17:24  Type: P  Check: 7067  ..

       1  XDKGSYLPRP TPPRPIYNRX

!!AA_SEQUENCE 1.0
ID   AAY72450 standard; peptide; 21 AA.
XX
AC   AAY72450;
XX
DT   06-AUG-2003  (revised)
DT   24-APR-2001  (first entry)
XX
DE   Pyrrhocoricin-modified peptide 18.
XX
KW   Pyrrhocoricin-derived peptide; antibacterial; fungicidal; therapy;
KW   fungal infection; bacterial infection; candidiasis; drug development.
XX
OS   Pyrrhocoris apterus.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   1
FT                   /note= "N-terminal biotin"
XX
PN   WO200078956-A1.
XX
PD   28-DEC-2000.
XX
PF   21-JUN-2000; 2000WO-US016989.
XX
PR   23-JUN-1999;   99US-0140606P.
PR   15-SEP-1999;   99US-0154135P.
XX
PA   (WIST-) WISTAR INST ANATOMY & BIOLOGY.
XX
PI   Otvos L;
XX
DR   WPI; 2001-112323/12.
XX
PT   Polypeptides derived from the peptide pyrrhocoricin, useful for treating
PT   fungal infections and Gram negative/positive bacterial infections.
XX
PS   Claim 31; Page 46; 75pp; English.
XX
CC   The present peptide sequence is active Pyrrhocoricin-modified Peptide 18.
CC   Pyrrhocoricin is a glycopeptide characterised by the presence of a
CC   disaccharide in the mid-chain position. The invention relates to
CC   pyrrhocoricin-derived peptides which have anti-bacterial or anti-fungal
CC   activity. These peptides have metabolic stability in mammalian serum. The
CC   pyrrhocoricin-derived peptides are used in the treatment of bacterial
CC   infections caused by Gram positive or Gram negative bacterium and fungal
CC   infections of skin, nails, mucus membranes and intestines e.g.,
CC   candidiasis. These peptides are also useful in anti-bacterial or anti-
CC   fungal pharmaceutical compositions, drug development and identification
CC   of other antibiotic or anti-fungal compounds. (Updated on 06-AUG-2003 to
CC   correct OS field.)
XX
SQ   Sequence 21 AA;

  AAY72450  Length: 21  March 11, 2004 17:24  Type: P  Check: 8536  ..

       1  KVDKGSYLPR PTPPRPIYNR N
```

```
!!AA_SEQUENCE 1.0
ID   AAY72456 standard; peptide; 20 AA.
XX
AC   AAY72456;
XX
DT   06-AUG-2003  (revised)
DT   24-APR-2001  (first entry)
XX
DE   Pyrrhocoricin-modified Peptide 24.
XX
KW   Pyrrhocoricin-derived peptide; antibacterial; fungicidal; therapy;
KW   fungal infection; bacterial infection; candidiasis; drug development.
XX
OS   Pyrrhocoris apterus.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   20
FT                   /note= "Beta-acetyl-2,3-diamino propionic acid"
XX
PN   WO200078956-A1.
XX
PD   28-DEC-2000.
XX
PF   21-JUN-2000; 2000WO-US016989.
XX
PR   23-JUN-1999;   99US-0140606P.
PR   15-SEP-1999;   99US-0154135P.
XX
PA   (WIST-) WISTAR INST ANATOMY & BIOLOGY.
XX
PI   Otvos L;
XX
DR   WPI; 2001-112323/12.
XX
PT   Polypeptides derived from the peptide pyrrhocoricin, useful for treating
PT   fungal infections and Gram negative/positive bacterial infections.
XX
PS   Claim 36; Page 47; 75pp; English.
XX
CC   The present peptide sequence is active Pyrrhocoricin-modified Peptide 24.
CC   Pyrrhocoricin is a glycopeptide characterised by the presence of a
CC   disaccharide in the mid-chain position. The invention relates to
CC   pyrrhocoricin-derived peptides which have anti-bacterial or anti-fungal
CC   activity. These peptides have metabolic stability in mammalian serum. The
CC   pyrrhocoricin-derived peptides are used in the treatment of bacterial
CC   infections caused by Gram positive or Gram negative bacterium and fungal
CC   infections of skin, nails, mucus membranes and intestines e.g.,
CC   candidiasis. These peptides are also useful in anti-bacterial or anti-
CC   fungal pharmaceutical compositions, drug development and identification
CC   of other antibiotic or anti-fungal compounds. (Updated on 06-AUG-2003 to
CC   correct OS field.)
XX
SQ   Sequence 20 AA;

  AAY72456  Length: 20  March 11, 2004 17:24  Type: P  Check: 7065  ..

       1  VDKGSYLPRP TPPRPIYNRX

!!AA_SEQUENCE 1.0
ID   AAY72445 standard; peptide; 21 AA.
XX
AC   AAY72445;
XX
DT   06-AUG-2003  (revised)
DT   24-APR-2001  (first entry)
XX
DE   Pyrrhocoricin-modified Peptide 10.
XX
KW   Pyrrhocoricin-derived peptide; antibacterial; fungicidal; therapy;
KW   fungal infection; bacterial infection; candidiasis; drug development.
XX
OS   Pyrrhocoris apterus.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   1
FT                   /note= "N-terminal acetyl"
FT   Modified-site   21
FT                   /note= "Modified with 2-acetamido-2-deoxyglucose
FT                   (GlcNAc)"
XX
PN   WO200078956-A1.
XX
PD   28-DEC-2000.
XX
PF   21-JUN-2000; 2000WO-US016989.
XX
PR   23-JUN-1999;   99US-0140606P.
PR   15-SEP-1999;   99US-0154135P.
XX
PA   (WIST-) WISTAR INST ANATOMY & BIOLOGY.
XX
PI   Otvos L;
XX
DR   WPI; 2001-112323/12.
XX
PT   Polypeptides derived from the peptide pyrrhocoricin, useful for treating
PT   fungal infections and Gram negative/positive bacterial infections.
XX
PS   Claim 28; Page 46; 75pp; English.
XX
CC   The present peptide sequence is active Pyrrhocoricin-modified Peptide 10.
CC   Pyrrhocoricin is a glycopeptide characterised by the presence of a
CC   disaccharide in the mid-chain position. The invention relates to
CC   pyrrhocoricin-derived peptides which have anti-bacterial or anti-fungal
CC   activity. These peptides have metabolic stability in mammalian serum. The
CC   pyrrhocoricin-derived peptides are used in the treatment of bacterial
CC   infections caused by Gram positive or Gram negative bacterium and fungal
CC   infections of skin, nails, mucus membranes and intestines e.g.,
CC   candidiasis. These peptides are also useful in anti-bacterial or anti-
CC   fungal pharmaceutical compositions, drug development and identification
CC   of other antibiotic or anti-fungal compounds. (Updated on 06-AUG-2003 to
CC   correct OS field.)
XX
SQ   Sequence 21 AA;

  AAY72445  Length: 21  March 11, 2004 17:24  Type: P  Check: 8536  ..

       1  KVDKGSYLPR PTPPRPIYNR N

!!AA_SEQUENCE 1.0
ID   AAY72437 standard; peptide; 20 AA.
XX
AC   AAY72437;
XX
DT   06-AUG-2003  (revised)
DT   24-APR-2001  (first entry)
XX
DE   Pyrrhocoricin-modified Peptide 1.
XX
KW   Pyrrhocoricin-derived peptide; antibacterial; fungicidal; therapy;
KW   fungal infection; bacterial infection; candidiasis; drug development.
XX
OS   Pyrrhocoris apterus.
OS   Synthetic.
XX
PN   WO200078956-A1.
XX
PD   28-DEC-2000.
XX
PF   21-JUN-2000; 2000WO-US016989.
XX
PR   23-JUN-1999;   99US-0140606P.
PR   15-SEP-1999;   99US-0154135P.
```

XX
PA (WIST-) WISTAR INST ANATOMY & BIOLOGY.
XX
PI Otvos L;
XX
DR WPI; 2001-112323/12.
XX
PT Polypeptides derived from the peptide pyrrhocoricin, useful for treating
PT fungal infections and Gram negative/positive bacterial infections.
XX
PS Example 1; Page 23; 75pp; English.
XX
CC The present peptide sequence is active Pyrrhocoricin-modified Peptide 1
CC in which the naturally occurring mid-chain glycosylation is deleted.
CC Pyrrhocoricin is a glycopeptide characterised by the presence of a
CC disaccharide in the mid-chain position. The invention relates to
CC pyrrhocoricin-derived peptides which have anti-bacterial or anti-fungal
CC activity. These peptides have metabolic stability in mammalian serum. The
CC pyrrhocoricin-derived peptides are used in the treatment of bacterial
CC infections caused by Gram positive or Gram negative bacterium and fungal
CC infections of skin, nails, mucus membranes and intestines e.g.,
CC candidiasis. These peptides are also useful in anti-bacterial or anti-
CC fungal pharmaceutical compositions, drug development and identification
CC of other antibiotic or anti-fungal compounds. (Updated on 06-AUG-2003 to
CC correct OS field.)
XX
SQ Sequence 20 AA;

AAY72437 Length: 20 March 11, 2004 17:24 Type: P Check: 6865 ..

1 VDKGSYLPRP TPPRPIYNRN

!!AA_SEQUENCE 1.0
ID AAY72433 standard; peptide; 20 AA.
XX
AC AAY72433;
XX
DT 06-AUG-2003 (revised)
DT 24-APR-2001 (first entry)
XX
DE Native Pyrrhocoricin, Peptide 2.
XX
KW Pyrrhocoricin-derived peptide; antibacterial; fungicidal; therapy;
KW fungal infection; bacterial infection; candidiasis; drug development.
XX
OS Pyrrhocoris apterus.
XX
FH Key Location/Qualifiers
FT Cleavage-site 5. .6
FT /label= Endopeptidase_cleavage_site
FT Modified-site 11
FT /note= "Modified with Galactose-2-acetamido-2- deoxy-
FT galactose (Gal-GalNAc)"
FT Cleavage-site 18. .19
FT /label= Endopeptidase_cleavage_site
XX
PN WO200078956-A1.
XX
PD 28-DEC-2000.
XX
PF 21-JUN-2000; 2000WO-US016989.
XX
PR 23-JUN-1999; 99US-0140606P.
PR 15-SEP-1999; 99US-0154135P.
XX
PA (WIST-) WISTAR INST ANATOMY & BIOLOGY.
XX
PI Otvos L;
XX
DR WPI; 2001-112323/12.
XX
PT Polypeptides derived from the peptide pyrrhocoricin, useful for treating
PT fungal infections and Gram negative/positive bacterial infections.
XX
PS Example 1; Page 23; 75pp; English.
XX
CC The present sequence is native pyrrhocoricin, Peptide 2 which is
CC glycosylated. Pyrrhocoricin is a glycopeptide characterised by the
CC presence of a disaccharide in the mid-chain position. The invention
CC relates to pyrrhocoricin-derived peptides which have anti-bacterial or
CC anti-fungal activity. These peptides have metabolic stability in
CC mammalian serum. The pyrrhocoricin-derived peptides are used in the
CC treatment of bacterial infections caused by Gram positive or Gram
CC negative bacterium and fungal infections of skin, nails, mucus membranes
CC and intestines e.g., candidiasis. These peptides are also useful in anti-
CC bacterial or anti-fungal pharmaceutical compositions, drug development
CC and identification of other antibiotic or anti-fungal compounds. (Updated
CC on 06-AUG-2003 to correct OS field.)
XX
SQ Sequence 20 AA;

AAY72433 Length: 20 March 11, 2004 17:24 Type: P Check: 6865 ..

1 VDKGSYLPRP TPPRPIYNRN

!!AA_SEQUENCE 1.0
ID AAY72435 standard; peptide; 20 AA.
XX
AC AAY72435;
XX
DT 06-AUG-2003 (revised)
DT 24-APR-2001 (first entry)
XX
DE Pyrrhocoricin-modified peptide #1 for multi-peptide construction.
XX
KW Pyrrhocoricin-derived peptide; antibacterial; fungicidal; therapy;
KW fungal infection; bacterial infection; candidiasis; drug development.
XX
OS Pyrrhocoris apterus.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Modified-site 1
FT /note= "Homoproline or 1-aminocyclo-hexane carboxylic
FT acid"
FT Cross-links 20
FT /note= "The carboxy group of the 2-amino-3-acetylamino-
FT propionic acid residue 20 of AAY72498 is condensed onto
FT the side chain amino group of 2,3-diamino propionic acid
FT residue 20 of AAY72435 to cross link the two peptides
FT into a multipeptide"
FT Modified-site 20
FT /note= "2,3-diamino propionic acid amide"
XX
PN WO200078956-A1.
XX
PD 28-DEC-2000.
XX
PF 21-JUN-2000; 2000WO-US016989.
XX
PR 23-JUN-1999; 99US-0140606P.
PR 15-SEP-1999; 99US-0154135P.
XX
PA (WIST-) WISTAR INST ANATOMY & BIOLOGY.
XX
PI Otvos L;
XX
DR WPI; 2001-112323/12.
XX
PT Polypeptides derived from the peptide pyrrhocoricin, useful for treating
PT fungal infections and Gram negative/positive bacterial infections.
XX
PS Claim 51; Page 50; 75pp; English.
XX
CC The present peptide sequence is Pyrrhocoricin-modified peptide used for
CC multipeptide construction. Pyrrhocoricin is a glycopeptide characterised

CC by the presence of a disaccharide in the mid-chain position. The
CC invention relates to pyrrhocoricin-derived peptides which have anti-
CC bacterial or anti-fungal activity. These peptides have metabolic
CC stability in mammalian serum. The pyrrhocoricin-derived peptides are used
CC in the treatment of bacterial infections caused by Gram positive or Gram
CC negative bacterium and fungal infections of skin, nails, mucus membranes
CC and intestines e.g., candidiasis. These peptides are also useful in anti-
CC bacterial or anti-fungal pharmaceutical compositions, drug development
CC and identification of other antibiotic or anti-fungal compounds. (Updated
CC on 06-AUG-2003 to correct OS field.)
XX
SQ Sequence 20 AA;

AAY72435 Length: 20 March 11, 2004 17:24 Type: P Check: 7067 ..

1 XDKGSYLPRP TPPRPIYNRX

!!AA_SEQUENCE 1.0
ID AAY72438 standard; peptide; 24 AA.
XX
AC AAY72438;
XX
DT 06-AUG-2003 (revised)
DT 24-APR-2001 (first entry)
XX
DE Pyrrhocoricin-modified Peptide 3.
XX
KW Pyrrhocoricin-derived peptide; antibacterial; fungicidal; therapy;
KW fungal infection; bacterial infection; candidiasis; drug development.
XX
OS Pyrrhocoris apterus.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Modified-site 1
FT /note= "N-terminal acetyl"
XX
PN WO200078956-A1.
XX
PD 28-DEC-2000.
XX
PF 21-JUN-2000; 2000WO-US016989.
XX
PR 23-JUN-1999; 99US-0140606P.
PR 15-SEP-1999; 99US-0154135P.
XX
PA (WIST-) WISTAR INST ANATOMY & BIOLOGY.
XX
PI Otvos L;
XX
DR WPI; 2001-112323/12.
XX
PT Polypeptides derived from the peptide pyrrhocoricin, useful for treating
PT fungal infections and Gram negative/positive bacterial infections.
XX
PS Claim 22; Page 45; 75pp; English.
XX
CC The present peptide sequence is active Pyrrhocoricin-modified Peptide 3.
CC Pyrrhocoricin is a glycopeptide characterised by the presence of a
CC disaccharide in the mid-chain position. The invention relates to
CC pyrrhocoricin-derived peptides which have anti-bacterial or anti-fungal
CC activity. These peptides have metabolic stability in mammalian serum. The
CC pyrrhocoricin-derived peptides are used in the treatment of bacterial
CC infections caused by Gram positive or Gram negative bacterium and fungal
CC infections of skin, nails, mucus membranes and intestines e.g.,
CC candidiasis. These peptides are also useful in anti-bacterial or anti-
CC fungal pharmaceutical compositions, drug development and identification
CC of other antibiotic or anti-fungal compounds. (Updated on 06-AUG-2003 to
CC correct OS field.)
XX
SQ Sequence 24 AA;

AAY72438 Length: 24 March 11, 2004 17:24 Type: P Check: 4000 ..

1 KVDKVDKGSY LPRPTPPRPI YNRN

!!AA_SEQUENCE 1.0
ID AAY72446 standard; peptide; 21 AA.
XX
AC AAY72446;
XX
DT 06-AUG-2003 (revised)
DT 24-APR-2001 (first entry)
XX
DE Pyrrhocoricin-modified Peptide 11.
XX
KW Pyrrhocoricin-derived peptide; antibacterial; fungicidal; therapy;
KW fungal infection; bacterial infection; candidiasis; drug development.
XX
OS Pyrrhocoris apterus.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Modified-site 1
FT /note= "N-terminal acetyl"
FT Modified-site 21
FT /note= "Modified with triacetyl-2-acetamido-2-
FT deoxyglucose (Ac3-GlcNAc)"
XX
PN WO200078956-A1.
XX
PD 28-DEC-2000.
XX
PF 21-JUN-2000; 2000WO-US016989.
XX
PR 23-JUN-1999; 99US-0140606P.
PR 15-SEP-1999; 99US-0154135P.
XX
PA (WIST-) WISTAR INST ANATOMY & BIOLOGY.
XX
PI Otvos L;
XX
DR WPI; 2001-112323/12.
XX
PT Polypeptides derived from the peptide pyrrhocoricin, useful for treating
PT fungal infections and Gram negative/positive bacterial infections.
XX
PS Claim 29; Page 46; 75pp; English.
XX
CC The present peptide sequence is active Pyrrhocoricin-modified Peptide 11.
CC Pyrrhocoricin is a glycopeptide characterised by the presence of a
CC disaccharide in the mid-chain position. The invention relates to
CC pyrrhocoricin-derived peptides which have anti-bacterial or anti-fungal
CC activity. These peptides have metabolic stability in mammalian serum. The
CC pyrrhocoricin-derived peptides are used in the treatment of bacterial
CC infections caused by Gram positive or Gram negative bacterium and fungal
CC infections of skin, nails, mucus membranes and intestines e.g.,
CC candidiasis. These peptides are also useful in anti-bacterial or anti-
CC fungal pharmaceutical compositions, drug development and identification
CC of other antibiotic or anti-fungal compounds. (Updated on 06-AUG-2003 to
CC correct OS field.)
XX
SQ Sequence 21 AA;

AAY72446 Length: 21 March 11, 2004 17:24 Type: P Check: 8536 ..

1 KVDKGSYLPR PTPPRPIYNR N

!!AA_SEQUENCE 1.0
ID ABG73945 standard; peptide; 19 AA.
XX
AC ABG73945;
XX
DT 31-MAR-2003 (first entry)
XX
DE Cell wall/cell membrane transport peptide #4.

XX
KW Transport peptide; cell wall; cell membrane; protein nucleic acid; PNA;
KW genetically modified micro-organism; bacterial infection.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Modified-site 1
FT /label= OTHER
FT /note= "Lys is hydrogenated"
FT Modified-site 19
FT /label= OTHER
FT /note= "Cys is covalently linked via an smcc (not
FT defined) polyethylene-glycol moiety to the nucleic acid
FT sequence appearing as ABX15985"
XX
PN WO200279467-A2.
XX
PD 10-OCT-2002.
XX
PF 26-MAR-2002; 2002WO-DK000208.
XX
PR 29-MAR-2001; 2001DK-00000523.
XX
PA (UYKO-) UNIV KOBENHAVNS.
XX
PI Nielsen PE, Good L;
XX
DR WPI; 2003-103273/09.
XX
PT Selecting genetically modified cells useful for isolation and industrial
PT growth of transformed organisms comprises treating the modified cells
PT with an antisense or antigene construct directed against the essential
PT gene X of the cells.
XX
PS Claim 16; Page 51; 92pp; English.
XX
CC The invention relates to selecting genetically modified cells comprising:
CC (a) modifying cells containing a growth essential gene X, with a vector
CC containing gene Y; and (b) treating the modified cells with an antisense
CC or antigene construct directed against the essential gene X of the cells
CC to obtain preferential growth of the modified cells over other non-
CC modified cells. Also included is a product manufactured fully or
CC partially by use of the new method. The method is useful for selecting
CC genetically modified cells and manufacturing a product. It is useful for
CC research the isolation and industrial growth maintenance of transformed
CC organisms. The new method has the advantage of selecting and maintaining
CC a plasmid containing bacterial culture without the use of antibiotics.
CC This has a wide variety of applications in research, development, and
CC industrial production involving genetically modified micro-organisms. The
CC method inhibits bacterial infections in eukaryotic cell cultures. The
CC present sequence is a cell wall/cell membrane transport peptide which is
CC incorporated into a peptide nucleic acid (PNA) antisense molecules for
CC use in the method of the invention
XX
SQ Sequence 19 AA;

ABG73945 Length: 19 March 11, 2004 17:24 Type: P Check: 5020 ..

1 VDKGSYLPRP TPPRPIYNC

!!AA_SEQUENCE 1.0
ID ADD35367 standard; peptide; 20 AA.
XX
AC ADD35367;
XX
DT 15-JAN-2004 (first entry)
XX
DE Antimicrobial peptide pyrrhocoricin.
XX
KW antimicrobial; ophthalmic; prostaglandin; hypotensive; ophthalmological;
KW intraocular pressure; glaucoma; ocular hypertension; hyperaemia;
KW irritation; inflammation; conjunctiva; ocular cell dysplasia;
KW iridial melanocyte hyperplasia; hyperpigmentation.
XX
OS Unidentified.
XX
PN WO2003079997-A2.
XX
PD 02-OCT-2003.
XX
PF 21-MAR-2003; 2003WO-US008935.
XX
PR 21-MAR-2002; 2002US-0367071P.
XX
PA (CAYM-) CAYMAN CHEM CO.
XX
PI Maxey KM, Johnson J;
XX
DR WPI; 2004-011506/01.
XX
PT Ophthalmic solution useful for the treatment of increased intraocular
PT pressure comprises a prostaglandin of the F-series and an antimicrobial
PT peptide.
XX
PS Disclosure; Page 11; 11pp; English.
XX
CC The invention relates to a novel ophthalmic solution comprising a
CC prostaglandin of the F-series and an antimicrobial peptide. A solution of
CC the invention has hypotensive and ophthalmological activity. The solution
CC is useful for the treatment of increased intraocular pressure, such as
CC caused by glaucoma and for the reduction of ocular hypertension. The
CC prostaglandin and the antimicrobial peptide work synergistically, to
CC provide beneficial reduction in the incidence of irritant and toxic side
CC effects such as hyperaemia, irritation and inflammation of conjunctiva,
CC ocular cell dysplasia, iridial melanocyte hyperplasia, and
CC hyperpigmentation, associated with the prior art prostaglandin
CC compositions. The present sequence represents an antimicrobial peptide of
CC the invention.
XX
SQ Sequence 20 AA;

ADD35367 Length: 20 March 11, 2004 17:24 Type: P Check: 6865 ..

1 VDKGSYLPRP TPPRPIYNRN

```
! FINDPATTERNS on pir:* allowing 0 mismatches

!        1 DKGXXLPRPTPPRPIYXX                          March 11, 2004 16:54 ..


              S44465  ck: 6865  len: 20     ! pyrrhocoricin - Pyrrhocoris apterus

1                     DKGXXLPRPTPPRPIYXX
             2:     V DKGSYLPRPTPPRPIYNR N

 Databases searched:
        NBRF, Release 78.0, Released on 24Nov2003, Formatted on 25Nov2003

     Total finds:              1
    Total length:     96,191,526
 Total sequences:        283,366
        CPU time:          53.51
```

```
!!SEQUENCE_LIST 1.0
! FINDPATTERNS on pir:* allowing 0 mismatches

!        1 DKGXXLPRPTPPRPIYXX                              March 11, 2004 16:59 ..

PIR2:S44465           ck: 6865  len: 20     finds: 1     ! pyrrhocoricin - Pyrrhocoris ap

\\End of list

 Databases searched:
        NBRF, Release 78.0, Released on 24Nov2003, Formatted on 25Nov2003

     Total finds:              1
    Total length:     96,191,526
 Total sequences:        283,366
        CPU time:       01:26.44
```

```
!!AA_SEQUENCE 1.0
P1;S44465 - pyrrhocoricin - Pyrrhocoris apterus
C;Species: Pyrrhocoris apterus
C;Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 07-May-1999
C;Accession: S44465
R;Cociancich, S.; Dupont, A.; Hegy, G.; Lanot, R.; Holder, F.; Hetru, C.;
 Hoffmann, J.A.; Bulet, P.
Biochem. J. 300, 567-575, 1994
A;Title: Novel inducible antibacterial peptides from a hemipteran insect, the
 sap-sucking bug Pyrrhocoris apterus.
A;Reference number: S44463; MUID:94271176; PMID:8002963
A;Accession: S44465
A;Molecule type: protein
A;Residues: 1-20 <COC>
C;Function:
A;Description: antibacterial protein
A;Note: active against Gram-negative bacteria
C;Keywords: antibacterial; hemolymph; immune response

  S44465  Length: 20  March 11, 2004 17:23  Type: P  Check: 6865  ..

      1  VDKGSYLPRP TPPRPIYNRN
```

```
! FINDPATTERNS on swp:* allowing 0 mismatches

!        1 DKGXXLPRPTPPRPIYXX                          March 11, 2004 16:56 ..


          PYRR_PYRAP  ck: 6865  len: 20    ! P37362 pyrrhocoris apterus (sap sucking bug

1                    DKGXXLPRPTPPRPIYXX
             2:     V DKGSYLPRPTPPRPIYNR N

 Databases searched:
        SWISS-PROT, Release 42.7, Released on 15Dec2003, Formatted on 15Dec2003
        SPTREMBL, Release 25.0, Released on 17Oct2003, Formatted on 18Oct2003

     Total finds:              1
    Total length:    367,588,357
 Total sequences:      1,158,722
        CPU time:       04:03.20
```

```
!!SEQUENCE_LIST 1.0
! FINDPATTERNS on swp:* allowing 0 mismatches

!        1 DKGXXLPRPTPPRPIYXX                          March 11, 2004 17:01 ..


SW:PYRR_PYRAP         ck: 6865  len: 20     finds: 1    ! P37362 pyrrhocoris apterus (sa

\\End of list

 Databases searched:
        SWISS-PROT, Release 42.7, Released on 15Dec2003, Formatted on 15Dec2003
        SPTREMBL, Release 25.0, Released on 17Oct2003, Formatted on 18Oct2003

     Total finds:              1
    Total length:    367,588,357
 Total sequences:      1,158,722
        CPU time:         06:13.80
```

```
!!AA_SEQUENCE 1.0
ID   PYRR_PYRAP     STANDARD;      PRT;    20 AA.
AC   P37362; P80307;
DT   01-OCT-1994 (Rel. 30, Created)
DT   01-OCT-1994 (Rel. 30, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Pyrrhocoricin.
OS   Pyrrhocoris apterus (Sap sucking bug).
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Paraneoptera; Hemiptera; Euhemiptera; Heteroptera;
OC   Panheteroptera; Pentatomomorpha; Pyrrhocoroidea; Pyrrhocoridae;
OC   Pyrrhocoris.
OX   NCBI_TaxID=37000;
RN   [1]
RP   SEQUENCE.
RC   TISSUE=Hemolymph;
RX   MEDLINE=94271176; PubMed=8002963;
RA   Cociancich S., Dupont A., Hegy G., Lanot R., Holder F., Hetru C.,
RA   Hoffmann J.A., Bulet P.;
RT   "Novel inducible antibacterial peptides from a hemipteran insect, the
RT   sap-sucking bug Pyrrhocoris apterus.";
RL   Biochem. J. 300:567-575(1994).
RN   [2]
RP   CARBOHYDRATE-LINKAGE SITE THR-11.
RX   MEDLINE=99177428; PubMed=10076062;
RA   Hoffmann R., Bulet P., Urge L., Otvoes L. Jr.;
RT   "Range of activity and metabolic stability of synthetic antibacterial
RT   glycopeptides from insects.";
RL   Biochim. Biophys. Acta 1426:459-467(1999).
CC   -!- FUNCTION: Antibacterial peptide. Affects Gram-negative bacteria
CC       E.coli 1106, P.aeruginosa, E.coli D22 and E.cloacae and
CC       Gram-positive bacteria M.luteus and B.subtilis.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- PTM: O-LINKED GLYCAN CONSISTS OF A GAL-GALNAC DISACCHARIDE, O-
CC       GLYCOSYLATION IS ESSENTIAL FOR FULL BIOLOGICAL ACTIVITY.
CC   -!- SIMILARITY: TO DROSOPHILA DROSOCIN.
DR   PIR; S44465; S44465.
KW   Antibiotic; Glycoprotein; Insect immunity; Hemolymph.
FT   CARBOHYD     11     11       O-LINKED (GALNAC. . .).
SQ   SEQUENCE   20 AA;  2341 MW;  F4320EC2FF29462C CRC64;

 PYRR_PYRAP  Length: 20  March 11, 2004 17:23  Type: P  Check: 6865  ..

       1  VDKGSYLPRP TPPRPIYNRN
```

```
> O <
O|  |O  IntelliGenetics
> O <

Quest - Quick User-directed Expression Search Tool
Release 5.4

                -- Outline of search "seq1-iss" --

Selected search type is key against sequence data banks or files.
Selected scope is Sequence.
Selected sequence key from "new.key":
  seq1 (AA)
1       followed by
2         dkg
2         any character
2         any character
2         lprptpprpiy
2         any character
2         any character

Selected data banks and files:

   Data bank : Issued_AA , all entries

                    -- Output Parameters --

Format Options:                        File Options:
  Nucleic acid code matching   Exact     Indirect file               No
  Find non-matching hits only  No        Sequence or key file        No
  Report key used              Yes       List of hits                No
  Note position of hit         Yes       Hit display                 Yes
  Display full annotations     Yes       Name and annotations        Yes
  Sequence context             50


                      -- Run Parameters --

  Run mode                     Interactive
    Prompt on hit              Yes
    Bell on hit                No


No hits found.


                    -- Search Statistics --

Times:                  CPU                 Total Elapsed
                    00:01:47.05             00:01:49.00

Number of sequences searched:                  389531
Number of sequence hits:                            0
Number of separate matches:                         0
Number of sequence hits saved:                      0
```

```
!!SEQUENCE_LIST 1.0
! FINDPATTERNS on geneseqp:* allowing 0 mismatches

!        1 DKGXXLPRPTPPRPIYXX                          March 11, 2004 17:15 ..

GENESEQP1990S:AAR50300  ck: 6865  len: 20     finds: 1    ! Aar50300 Anti-bacterial glyc
GENESEQP2001S:AAG62740  ck: 4080  len: 18     finds: 1    ! Aag62740 Amino acid sequence
GENESEQP2001S:AAG62734  ck: 6865  len: 20     finds: 1    ! Aag62734 Amino acid sequence
GENESEQP2001S:AAG62743  ck: 8536  len: 21     finds: 1    ! Aag62743 Amino acid sequence
GENESEQP2001S:AAG62756  ck: 8536  len: 21     finds: 1    ! Aag62756 Amino acid sequence
GENESEQP2001S:AAY72457  ck: 6865  len: 20     finds: 1    ! Aay72457 Pyrrhocoricin-modif
GENESEQP2001S:AAY72439  ck: 8543  len: 21     finds: 1    ! Aay72439 Pyrrhocoricin-modif
GENESEQP2001S:AAY72444  ck: 8746  len: 21     finds: 1    ! Aay72444 Pyrrhocoricin-modif
GENESEQP2001S:AAY72454  ck: 8753  len: 21     finds: 1    ! Aay72454 Pyrrhocoricin-modif
GENESEQP2001S:AAY72455  ck: 6863  len: 20     finds: 1    ! Aay72455 Pyrrhocoricin-modif
GENESEQP2001S:AAY72461  ck: 2100  len: 23     finds: 1    ! Aay72461 Pyrrhocoricin-modif
GENESEQP2001S:AAY72442  ck: 6865  len: 20     finds: 1    ! Aay72442 Pyrrhocoricin-modif
GENESEQP2001S:AAY72448  ck: 8326  len: 21     finds: 1    ! Aay72448 Pyrrhocoricin-modif
GENESEQP2001S:AAY72449  ck: 4782  len: 29     finds: 1    ! Aay72449 Pyrrhocoricin-modif
GENESEQP2001S:AAY72424  ck: 4080  len: 18     finds: 1    ! Aay72424 Pyrrhocoricin based
GENESEQP2001S:AAY72440  ck: 8536  len: 21     finds: 1    ! Aay72440 Pyrrhocoricin-modif
GENESEQP2001S:AAY72441  ck: 6867  len: 20     finds: 1    ! Aay72441 Pyrrhocoricin-modif
GENESEQP2001S:AAY72443  ck: 6898  len: 20     finds: 1    ! Aay72443 Pyrrhocoricin-modif
GENESEQP2001S:AAY72447  ck: 6865  len: 20     finds: 1    ! Aay72447 Pyrrhocoricin-modif
GENESEQP2001S:AAY72453  ck: 7067  len: 20     finds: 1    ! Aay72453 Pyrrhocoricin-modif
GENESEQP2001S:AAY72451  ck: 8536  len: 21     finds: 1    ! Aay72451 Pyrrhocoricin-modif
GENESEQP2001S:AAY72452  ck: 8326  len: 21     finds: 1    ! Aay72452 Pyrrhocoricin-modif
GENESEQP2001S:AAY72498  ck: 7067  len: 20     finds: 1    ! Aay72498 Pyrrhocoricin-modif
GENESEQP2001S:AAY72450  ck: 8536  len: 21     finds: 1    ! Aay72450 Pyrrhocoricin-modif
GENESEQP2001S:AAY72456  ck: 7065  len: 20     finds: 1    ! Aay72456 Pyrrhocoricin-modif
GENESEQP2001S:AAY72445  ck: 8536  len: 21     finds: 1    ! Aay72445 Pyrrhocoricin-modif
GENESEQP2001S:AAY72437  ck: 6865  len: 20     finds: 1    ! Aay72437 Pyrrhocoricin-modif
GENESEQP2001S:AAY72433  ck: 6865  len: 20     finds: 1    ! Aay72433 Native Pyrrhocorici
GENESEQP2001S:AAY72435  ck: 7067  len: 20     finds: 1    ! Aay72435 Pyrrhocoricin-modif
GENESEQP2001S:AAY72438  ck: 4000  len: 24     finds: 1    ! Aay72438 Pyrrhocoricin-modif
GENESEQP2001S:AAY72446  ck: 8536  len: 21     finds: 1    ! Aay72446 Pyrrhocoricin-modif
GENESEQP2003AS:ABG73945  ck: 5020  len: 19     finds: 1    ! Abg73945 Cell wall/cell mem
GENESEQP2004S:ADD35367  ck: 6865  len: 20     finds: 1    ! Add35367 Antimicrobial pepti
\\End of list
```

```
Databases searched:
      EMBL, Release 2.0, Released on 29Jan2004, Formatted on 12Feb2004

    Total finds:              33
   Total length:     282,547,505
Total sequences:       1,586,107
       CPU time:        07:14.61
```

```
1                    DKGXXLPRPTPPRPIYXX
              2:     X DKGSYLPRPTPPRPIYNR X

             AAY72438  ck: 4000  len: 24    ! Aay72438 Pyrrhocoricin-modified Peptide 3.

1                    DKGXXLPRPTPPRPIYXX
              6: KVDKV DKGSYLPRPTPPRPIYNR N

             AAY72446  ck: 8536  len: 21    ! Aay72446 Pyrrhocoricin-modified Peptide 11.

1                    DKGXXLPRPTPPRPIYXX
              3:    KV DKGSYLPRPTPPRPIYNR N

             ABG73945  ck: 5020  len: 19    ! Abg73945 Cell wall/cell membrane transport

1                    DKGXXLPRPTPPRPIYXX
              2:     V DKGSYLPRPTPPRPIYNC

             ADD35367  ck: 6865  len: 20    ! Add35367 Antimicrobial peptide pyrrhocorici

1                    DKGXXLPRPTPPRPIYXX
              2:     V DKGSYLPRPTPPRPIYNR N

 Databases searched:
        EMBL, Release 2.0, Released on 29Jan2004, Formatted on 12Feb2004

     Total finds:             33
    Total length:    282,547,505
 Total sequences:      1,586,107
        CPU time:       04:23.08
```

```
! FINDPATTERNS on geneseqp:* allowing 0 mismatches
!       1 DKGXXLPRPTPPRPIYXX                          March 11, 2004 17:09 ..

          AAR50300  ck: 6865  len: 20    ! Aar50300 Anti-bacterial glycopeptide #9 ind
1                   DKGXXLPRPTPPRPIYXX
 2:      V DKGSYLPRPTPPRPIYNR N

          AAG62740  ck: 4080  len: 18    ! Aag62740 Amino acid sequence of modified ar
1                   DKGXXLPRPTPPRPIYXX
 1:        DKGXXLPRPTPPRPIYXX

          AAG62734  ck: 6865  len: 20    ! Aag62734 Amino acid sequence of antibacteri
1                   DKGXXLPRPTPPRPIYXX
 2:      V DKGSYLPRPTPPRPIYNR N

          AAG62743  ck: 8536  len: 21    ! Aag62743 Amino acid sequence of modified ar
1                   DKGXXLPRPTPPRPIYXX
 3:     KV DKGSYLPRPTPPRPIYNR N

          AAG62756  ck: 8536  len: 21    ! Aag62756 Amino acid sequence of modified ar
1                   DKGXXLPRPTPPRPIYXX
 3:     KV DKGSYLPRPTPPRPIYNR N

          AAY72457  ck: 6865  len: 20    ! Aay72457 Pyrrhocoricin-modified Peptide 13.
1                   DKGXXLPRPTPPRPIYXX
 2:      V DKGSYLPRPTPPRPIYNR N

          AAY72439  ck: 8543  len: 21    ! Aay72439 Pyrrhocoricin-modified Peptide 4.
1                   DKGXXLPRPTPPRPIYXX
 3:     RV DKGSYLPRPTPPRPIYNR N

          AAY72444  ck: 8746  len: 21    ! Aay72444 Pyrrhocoricin-modified Peptide 9.
1                   DKGXXLPRPTPPRPIYXX
 3:     KV DKGSYLPRPTPPRPIYNR X

          AAY72454  ck: 8753  len: 21    ! Aay72454 Pyrrhocoricin-modified Peptide 22.
1                   DKGXXLPRPTPPRPIYXX
 3:     RV DKGSYLPRPTPPRPIYNR X

          AAY72455  ck: 6863  len: 20    ! Aay72455 Pyrrhocoricin-modified Peptide 23.
1                   DKGXXLPRPTPPRPIYXX
 2:      X DKGAFLPRPTPPRPIYNR X

          AAY72461  ck: 2100  len: 23    ! Aay72461 Pyrrhocoricin-modified peptide. 8/
1                   DKGXXLPRPTPPRPIYXX
 5:   VDKV DKGSYLPRPTPPRPIYNR N

          AAY72442  ck: 6865  len: 20    ! Aay72442 Pyrrhocoricin-modified Peptide 7.
1                   DKGXXLPRPTPPRPIYXX
 2:      V DKGSYLPRPTPPRPIYNR N

          AAY72448  ck: 8326  len: 21    ! Aay72448 Pyrrhocoricin-modified Peptide 16.
1                   DKGXXLPRPTPPRPIYXX
 3:     KV DKGSYLPRPTPPRPIYNR D

          AAY72449  ck: 4782  len: 29    ! Aay72449 Pyrrhocoricin-modified Peptide 17.
1                   DKGXXLPRPTPPRPIYXX
11: RPLKV DKGSYLPRPTPPRPIYNR N

          AAY72424  ck: 4080  len: 18    ! Aay72424 Pyrrhocoricin based generic pepti
1                   DKGXXLPRPTPPRPIYXX
 1:        DKGXXLPRPTPPRPIYXX

          AAY72440  ck: 8536  len: 21    ! Aay72440 Pyrrhocoricin-modified Peptide 5.
1                   DKGXXLPRPTPPRPIYXX
 3:     KV DKGSYLPRPTPPRPIYNR N

          AAY72441  ck: 6867  len: 20    ! Aay72441 Pyrrhocoricin-modified Peptide 6.
1                   DKGXXLPRPTPPRPIYXX
 2:      X DKGSYLPRPTPPRPIYNR N

          AAY72443  ck: 6898  len: 20    ! Aay72443 Pyrrhocoricin-modified Peptide 8.
1                   DKGXXLPRPTPPRPIYXX
 3:     KV DKGSYLPRPTPPRPIYNR

          AAY72447  ck: 6865  len: 20    ! Aay72447 Pyrrhocoricin-modified Peptide 12
1                   DKGXXLPRPTPPRPIYXX
 2:      V DKGSYLPRPTPPRPIYNR N

          AAY72453  ck: 7067  len: 20    ! Aay72453 Pyrrhocoricin-modified Peptide 21
1                   DKGXXLPRPTPPRPIYXX
 2:      X DKGSYLPRPTPPRPIYNR X

          AAY72451  ck: 8536  len: 21    ! Aay72451 Pyrrhocoricin-modified Peptide 19
1                   DKGXXLPRPTPPRPIYXX
 3:     KV DKGSYLPRPTPPRPIYNR N

          AAY72452  ck: 8326  len: 21    ! Aay72452 Pyrrhocoricin-modified Peptide 20
1                   DKGXXLPRPTPPRPIYXX
 3:     KV DKGSYLPRPTPPRPIYNR D

          AAY72498  ck: 7067  len: 20    ! Aay72498 Pyrrhocoricin-modified peptide #2
1                   DKGXXLPRPTPPRPIYXX
 2:      X DKGSYLPRPTPPRPIYNR X

          AAY72450  ck: 8536  len: 21    ! Aay72450 Pyrrhocoricin-modified peptide 18
1                   DKGXXLPRPTPPRPIYXX
 3:     KV DKGSYLPRPTPPRPIYNR N

          AAY72456  ck: 7065  len: 20    ! Aay72456 Pyrrhocoricin-modified Peptide 24
1                   DKGXXLPRPTPPRPIYXX
 2:      V DKGSYLPRPTPPRPIYNR X

          AAY72445  ck: 8536  len: 21    ! Aay72445 Pyrrhocoricin-modified Peptide 10
1                   DKGXXLPRPTPPRPIYXX
 3:     KV DKGSYLPRPTPPRPIYNR N

          AAY72437  ck: 6865  len: 20    ! Aay72437 Pyrrhocoricin-modified Peptide 1.
1                   DKGXXLPRPTPPRPIYXX
 2:      V DKGSYLPRPTPPRPIYNR N

          AAY72433  ck: 6865  len: 20    ! Aay72433 Native Pyrrhocoricin, Peptide 2.
1                   DKGXXLPRPTPPRPIYXX
 2:      V DKGSYLPRPTPPRPIYNR N

          AAY72435  ck: 7067  len: 20    ! Aay72435 Pyrrhocoricin-modified peptide #1
```

```
!!AA_SEQUENCE 1.0
ID   AAR50300 standard; peptide; 20 AA.
XX
AC   AAR50300;
XX
DT   25-MAR-2003  (revised)
DT   10-OCT-1994  (first entry)
XX
DE   Anti-bacterial glycopeptide #9 induced in Pyrrhocoris apterus.
XX
KW   Antibacterial glycopeptide; Diptera; septicaemia; Gram positive bacteria;
KW   Gram negative bacteria.
XX
OS   Pyrrhocoris apterus.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   11
FT                   /label= O-glycosylated
XX
PN   WO9405787-A1.
XX
PD   17-MAR-1994.
XX
PF   06-SEP-1993;   93WO-FR000853.
XX
PR   04-SEP-1992;   92FR-00010608.
XX
PA   (CNRS ) CNRS CENT NAT RECH SCI.
XX
PI   Bulet P,  Hetru C,  Dimarcq J,  Hoffmann J,  Van Dorsselaer A;
XX
DR   WPI; 1994-101192/12.
XX
PT   New antibacterial glycopeptide(s) derived from insects - for control of
PT   Gram negative and positive bacteria in human and veterinary medicine,
PT   agriculture, etc.
XX
PS   Claim 17; Page 9-10; 45pp; French.
XX
CC   This is a preferred example of an anti-bacterial glycopeptide induced in
CC   arthropods (esp. larval or adult insects) by injection of bacteria, a
CC   septic wound or other injury. The peptides contain at least one O-
CC   glycosylated residue and are useful for treatment of e.g. septicaemia,
CC   for oral or dental use and in gynaecology. (Updated on 25-MAR-2003 to
CC   correct PN field.)
XX
SQ   Sequence 20 AA;

  AAR50300  Length: 20  March 11, 2004 17:24  Type: P  Check: 6865  ..

       1  VDKGSYLPRP TPPRPIYNRN

!!AA_SEQUENCE 1.0
ID   AAG62740 standard; peptide; 18 AA.
XX
AC   AAG62740;
XX
DT   17-SEP-2001  (first entry)
XX
DE   Amino acid sequence of modified antibacterial peptide pyrrhocoricin.
XX
KW   Multi-helical lid; heat shock protein; hsp; protein folding;
KW   pathogenic infection; bacterial infection; antibacterial.
XX
OS   Unidentified.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   1
FT                   /note= "a moiety having a net positive charge is
FT                   attached"
XX
PN   WO200153509-A2.
XX
PD   26-JUL-2001.
XX
PF   19-JAN-2001; 2001WO-US001812.
XX
PR   21-JAN-2000; 2000US-0177565P.
PR   03-OCT-2000; 2000US-0237599P.
XX
PA   (WIST-) WISTAR INST ANATOMY & BIOLOGY.
PA   (UYCR-) UNIV CREIGHTON.
XX
PI   Otvos L,  Blaszczyk-Thurin M,  Rogers M,  Lovas S;
XX
DR   WPI; 2001-451911/48.
XX
PT   Composition, used to treat a pathogenic infection and eliminate a plant,
PT   insect, or animal pest, comprises a molecule that binds to a heat shock
PT   protein.
XX
PS   Disclosure; Page 111; 124pp; English.
XX
CC   The specification describes a composition that comprises a synthetic non-
CC   naturally occurring molecule that binds to a selected multi-helical lid
CC   of a heat shock protein (hsp) of a selected organism, where the molecule
CC   inhibits protein folding activity of the hsp, and a carrier, where
CC   exposure of the organism to the composition retards the growth and
CC   reproduction of the organism. The composition is used to treat a mammal
CC   suffering from a pathogenic infection, in the manufacture of a medicament
CC   for treating a mammal for a pathogenic infection, and to eliminate a
CC   plant, insect, or animal pest. It is used in the manufacture of a
CC   medicament for treating mammalian bacterial infection. The present
CC   sequence represents a modified antibacterial peptide, which may be used
CC   to produce the composition of the invention
XX
SQ   Sequence 18 AA;

  AAG62740  Length: 18  March 11, 2004 17:24  Type: P  Check: 4080  ..

       1  DKGXXLPRPT PPRPIYXX

!!AA_SEQUENCE 1.0
ID   AAG62734 standard; peptide; 20 AA.
XX
AC   AAG62734;
XX
DT   17-SEP-2001  (first entry)
XX
DE   Amino acid sequence of antibacterial peptide pyrrhocoricin.
XX
KW   Multi-helical lid; heat shock protein; hsp; protein folding;
KW   pathogenic infection; bacterial infection; antibacterial.
XX
OS   Unidentified.
XX
PN   WO200153509-A2.
XX
PD   26-JUL-2001.
XX
PF   19-JAN-2001; 2001WO-US001812.
XX
PR   21-JAN-2000; 2000US-0177565P.
PR   03-OCT-2000; 2000US-0237599P.
XX
PA   (WIST-) WISTAR INST ANATOMY & BIOLOGY.
PA   (UYCR-) UNIV CREIGHTON.
XX
PI   Otvos L,  Blaszczyk-Thurin M,  Rogers M,  Lovas S;
XX
DR   WPI; 2001-451911/48.
XX
PT   Composition, used to treat a pathogenic infection and eliminate a plant,
PT   insect, or animal pest, comprises a molecule that binds to a heat shock
PT   protein.
XX
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: March 11, 2004, 16:53:29 ; Search time 21 Seconds
(without alignments)
82.450 Million cell updates/sec

Title: US-09-980-804-1
Perfect score: 86
Sequence: 1 DKGXXLPRPTPPRPIYXX 18

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_78:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	81	94.2	20	2	S44465	pyrrhocoricin - Py
2	50	58.1	305	2	AB2149	hypothetical prote
3	48	55.8	1013	2	T46422	hypothetical prote
4	47	54.7	165	2	T45271	probable mini-circ
5	47	54.7	277	2	T49543	hypothetical prote
6	47	54.7	847	2	F96531	hypothetical prote
7	47	54.7	862	2	T46289	hypothetical prote
8	47	54.7	903	2	T00705	N-chimerin homolog
9	46	53.5	487	2	S42442	nuclear protein EB
10	46	53.5	1364	2	T00250	MEGF2 protein - hu
11	46	53.5	2774	2	A43359	microtubule-associ
12	45	52.3	242	2	S54156	extensin-like prot
13	45	52.3	1106	2	T31742	hypothetical prote
14	45	52.3	1255	2	T31065	diaphanous protein
15	45	52.3	1262	2	T13353	protein stn-B - fr
16	44	51.2	156	2	S75864	ribosomal protein
17	44	51.2	241	2	T26522	hypothetical prote
18	44	51.2	285	2	T30506	fibroblast growth
19	44	51.2	421	1	S11674	acrosin (EC 3.4.21
20	44	51.2	439	2	S51939	chitinase (EC 3.2.
21	44	51.2	691	2	A25704	synapsin I - rat
22	44	51.2	704	2	A30411	synapsin Ia - rat
23	44	51.2	705	2	A35363	synapsin I splice
24	44	51.2	706	2	E30411	synapsin Ia - bovi
25	44	51.2	882	2	S41034	hypothetical prote
26	44	51.2	899	2	B48586	suppressor of hair
27	44	51.2	2706	2	T28155	variant-specific s
28	43	50.0	131	2	B75277	hypothetical prote
29	43	50.0	165	2	A42361	DNA-directed RNA p
30	43	50.0	214	2	G75289	hypothetical prote
31	43	50.0	227	2	S50067	homeotic protein H
32	43	50.0	351	2	A56387	helix-loop-helix t
33	43	50.0	436	2	AH2447	molybdopterin bios
34	43	50.0	463	2	T01872	hypothetical prote
35	43	50.0	724	2	A38748	3-phosphatidylinos
36	43	50.0	724	2	A38749	3-phosphatidylinos
37	43	50.0	724	2	A38747	phosphatidlyinosit
38	43	50.0	760	2	T06291	extensin homolog T
39	43	50.0	925	2	T33732	probable excinucle
40	43	50.0	1119	2	T16720	hypothetical prote
41	43	50.0	1131	2	AD2005	RNA polymerase bet
42	43	50.0	1183	2	S63046	probable membrane
43	43	50.0	2133	2	T30637	hypothetical prote
44	43	50.0	3051	2	S42373	hypothetical prote
45	42	48.8	75	2	S14973	extensin class I (

ALIGNMENTS

RESULT 1
S44465
pyrrhocoricin - Pyrrhocoris apterus
C;Species: Pyrrhocoris apterus
C;Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 07-May-1999
C;Accession: S44465
R;Cociancich, S.; Dupont, A.; Hegy, G.; Lanot, R.; Holder, F.; Hetru, C.; Hoffmann, J.A
Biochem. J. 300, 567-575, 1994
A;Title: Novel inducible antibacterial peptides from a hemipteran insect, the sap-sucki
A;Reference number: S44463; MUID:94271176; PMID:8002963
A;Accession: S44465
A;Molecule type: protein
A;Residues: 1-20 <COC>
C;Function:
A;Description: antibacterial protein
A;Note: active against Gram-negative bacteria
C;Keywords: antibacterial; hemolymph; immune response

Query Match 94.2%; Score 81; DB 2; Length 20;
Best Local Similarity 87.5%; Pred. No. 3.3e-05;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy          1 DKGXXLPRPTPPRPIY 16
              |||  |||||||||||
Db          2 DKGSYLPRPTPPRPIY 17
```

RESULT 2
AB2149
hypothetical protein alr2745 [imported] - Nostoc sp. (strain PCC 7120)
C;Species: Nostoc sp. PCC 7120
A;Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C;Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Dec-2002
C;Accession: AB2149
R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguch
Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata,
DNA Res. 8, 205-213, 2001
A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium An
A;Reference number: AB1807; MUID:21595285; PMID:11759840
A;Accession: AB2149
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-305 <KUR>
A;Cross-references: GB:BA000019; PIDN:BAB74444.1; PID:g17131838; GSPDB:GN00179
A;Experimental source: strain PCC 7120
C;Genetics:
A;Gene: alr2745

Query Match 58.1%; Score 50; DB 2; Length 305;
Best Local Similarity 100.0%; Pred. No. 8.4;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          7 PRPTPPRP 14
              ||||||||
Db        187 PRPTPPRP 194
```

RESULT 3
T46422
hypothetical protein DKFZp434M2023.1 - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 18-Aug-2000
C;Accession: T46422
R;Blum, H.; Bauersachs, S.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
submitted to the Protein Sequence Database, January 2000
A;Reference number: Z23034
A;Accession: T46422
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-1013 <AAA>
A;Cross-references: EMBL:AL137480
A;Experimental source: adult testis; clone DKFZp434M2023
C;Genetics:
A;Note: DKFZp434M2023.1
C;Superfamily: WW repeat homology
F;210-248/Domain: WW repeat homology <WWR1>
F;591-629/Domain: WW repeat homology <WWR2>

  Query Match 55.8%; Score 48; DB 2; Length 1013;
  Best Local Similarity 66.7%; Pred. No. 53;
  Matches 8; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```
Qy          3 GXXLPRPTPPRP 14
              |  | ||||||
Db        161 GASAPPPTPPRP 172
```

RESULT 4
T45271
probable mini-circle protein [imported] - Streptomyces coelicolor (A3(2))
C;Species: Streptomyces coelicolor
A;Variety: A3(2)
C;Date: 31-Jan-2000 #sequence_revision 31-Jan-2000 #text_change 11-May-2000
C;Accession: T45271
R;Martinez-Costa, O.H.; Martin-Triana, A.J.; Martinez, E.; Fernandez-Moreno, M.A.; Malpa
J. Bacteriol. 181, 4353-4364, 1999
A;Title: An additinal regulatory gene for actinorhodin production in Streptomyces livida
A;Reference number: Z22953; MUID:99328982; PMID:10400594
A;Accession: T45271
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-165 <MAR>
A;Cross-references: EMBL:Y18817; PIDN:CAB51132.1
A;Experimental source: A3(2); strain J1501
C;Genetics:
A;Note: ORF7

  Query Match 54.7%; Score 47; DB 2; Length 165;
  Best Local Similarity 64.3%; Pred. No. 11;
  Matches 9; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

```
Qy          1 DKGXXLPRPTPPRP 14
              | |  |||| | ||
Db        117 DLGAPLPRPRPGRP 130
```

RESULT 5
T49543
hypothetical protein B21J21.220 [imported] - Neurospora crassa
C;Species: Neurospora crassa
C;Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 18-Aug-2000
C;Accession: T49543
R;Schulte, U.; Aign, V.; Hoheisel, J.; Brandt, P.; Fartmann, B.; Holland, R.; Nyakatura,
submitted to the Protein Sequence Database, May 2000
A;Reference number: Z25022
A;Accession: T49543
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-277 <SCH>
A;Cross-references: EMBL:AL355929; GSPDB:GN00116; NCSP:B21J21.220
A;Experimental source: BAC clone B21J21; strain OR74A
C;Genetics:
A;Gene: NCSP:B21J21.220
A;Map position: 6
A;Introns: 84/3; 99/2; 123/1
C;Superfamily: Neurospora crassa hypothetical protein B21J21.220

  Query Match 54.7%; Score 47; DB 2; Length 277;
  Best Local Similarity 100.0%; Pred. No. 19;
  Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          6 LPRPTPPR 13
              ||||||||
Db        228 LPRPTPPR 235
```

RESULT 6
F96531
hypothetical protein F13F21.7 [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C;Accession: F96531
R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonsc
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A;Reference number: A86141; MUID:21016719; PMID:11130712
A;Accession: F96531
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-847 <STO>
A;Cross-references: GB:AE005173; NID:g5430752; PIDN:AAD43152.1; GSPDB:GN00141
C;Genetics:
A;Gene: F13F21.7
A;Map position: 1

  Query Match 54.7%; Score 47; DB 2; Length 847;
  Best Local Similarity 63.6%; Pred. No. 60;
  Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

```
Qy          6 LPRPTPPRPIY 16
              :| |:|| |||
Db        540 MPSPSPPSPIY 550
```

RESULT 7
T46289
hypothetical protein DKFZp434A1010.1 - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 04-Feb-2000
C;Accession: T46289
R;Duesterhoeft, A.; Lauber, J.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
submitted to the Protein Sequence Database, January 2000
A;Reference number: Z23035
A;Accession: T46289
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-862 <AAA>
A;Cross-references: EMBL:AL137579
A;Experimental source: adult testis; clone DKFZp434A1010
C;Genetics:
A;Note: DKFZp434A1010.1

```
  Query Match             54.7%;  Score 47;  DB 2;  Length 862;
  Best Local Similarity   77.8%;  Pred. No. 62;
  Matches    7;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          8 RPTPPRPIY 16
              ||||| |:|
Db        736 RPTPPEPLY 744


RESULT 8
T00705
N-chimerin homolog F25965_3 - human
C;Species: Homo sapiens (man)
C;Date: 01-Feb-1999 #sequence_revision 01-Feb-1999 #text_change 05-Nov-1999
C;Accession: T00705
R;Lamerdin, J.E.; McCready, P.M.; Adamson, A.W.; Burkhart-Schultz, K.; Garcia, E.; Kyle,
hi, A.; Olsen, A.O.; Carrano, A.V.
submitted to the EMBL Data Library, October 1997
A;Description: Sequence analysis of a 1Mb region in 19q13.1.
A;Reference number: Z14199
A;Accession: T00705
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-903 <LAM>
A;Cross-references: EMBL:AC002398; NID:g2529398; PIDN:AAB81198.1; PID:g2477513
C;Genetics:
A;Map position: 19
A;Introns: 17/3; 68/3; 100/2; 148/3; 176/2; 212/2; 261/1; 312/2; 361/1; 513/1
A;Note: F25965_3

  Query Match             54.7%;  Score 47;  DB 2;  Length 903;
  Best Local Similarity   77.8%;  Pred. No. 64;
  Matches    7;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          8 RPTPPRPIY 16
              ||||| |:|
Db        777 RPTPPEPLY 785


RESULT 9
S42442
nuclear protein EBNA2 - human herpesvirus 4
C;Species: human herpesvirus 4, Epstein-Barr virus
C;Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 20-Jun-2000
C;Accession: S42442; S32988; S42447
R;Sample, J.; Hummel, M.; Braun, D.; Birkenbach, M.; Kieff, E.
Proc. Natl. Acad. Sci. U.S.A. 83, 5096-5100, 1986
A;Title: Nucleotide sequences of mRNAs encoding Epstein-Barr virus nuclear proteins: a p
A;Reference number: S42440; MUID:86259739; PMID:3460083
A;Accession: S42442
A;Molecule type: mRNA
A;Residues: 1-487 <SAM>
R;Farrell, P.J.
submitted to the EMBL Data Library, March 1988
A;Reference number: S32973
A;Accession: S32988
A;Molecule type: DNA
A;Residues: 1-487 <FAR>
A;Cross-references: EMBL:V01555; NID:g59074; PIDN:CAA24877.1; PID:g1632787
R;Dambaugh, T.; Hennessy, K.; Chamnankit, L.; Kieff, E.
Proc. Natl. Acad. Sci. U.S.A. 81, 7632-7636, 1984
A;Title: U2 region of Epstein-Barr virus DNA may encode Epstein-Barr nuclear antigen 2.
A;Reference number: S42447; MUID:85063846; PMID:6209719
A;Accession: S42447
A;Molecule type: DNA
A;Residues: 1-88,'PPP',89-487 <DAM>
A;Cross-references: EMBL:K03333; NID:g330443; PIDN:AAA45903.1; PID:g330444
C;Superfamily: hydroxyproline-rich glycoprotein

  Query Match             53.5%;  Score 46;  DB 2;  Length 487;
  Best Local Similarity   77.8%;  Pred. No. 47;
  Matches    7;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          7 PRPTPPRPI 15
              |||||| |:
Db        198 PRPTPPTPL 206


RESULT 10
T00250
MEGF2 protein - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 22-Jan-1999 #sequence_revision 22-Jan-1999 #text_change 21-Jul-2003
C;Accession: T00250
R;Nakayama, M.; Nakajima, D.; Nagase, T.; Nomura, N.; Seki, N.; Ohara, O.
Genomics 51, 27-34, 1998
A;Title: Identification of high-molecular-weight proteins with multiple EGF-like motifs
A;Reference number: Z14126; MUID:98360089; PMID:9693030
A;Accession: T00250
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-1364 <NAK>
A;Cross-references: EMBL:AB011536; NID:g3449297; PIDN:BAA32464.1; PID:g3449298
A;Experimental source: brain; clone HG1044
C;Genetics:
A;Gene: MEGF2
A;Map position: 3p21.2-p24.1
F;1-28/Domain: EGF homology (fragment) <EGF>
F;32-66/Domain: EGF homology <EGF1>
F;124-169/Domain: laminin-type EGF-like homology <LEG>

  Query Match             53.5%;  Score 46;  DB 2;  Length 1364;
  Best Local Similarity   61.5%;  Pred. No. 1.3e+02;
  Matches    8;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 DKGXXLPRPTPPR 13
              |:|  |||  |||
Db       1168 DRGSTLPRRQPPR 1180


RESULT 11
A43359
microtubule-associated protein MAP1A - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 13-Aug-1999
C;Accession: A43359; S22108
R;Langkopf, A.; Hammarback, J.A.; Muller, R.; Vallee, R.B.; Garner, C.C.
J. Biol. Chem. 267, 16561-16566, 1992
A;Title: Microtubule-associated proteins 1A and LC2. Two proteins encoded in one messen
A;Reference number: A43359; MUID:92355629; PMID:1379599
A;Accession: A43359
A;Molecule type: mRNA
A;Residues: 1-2774 <LAN>
A;Cross-references: GB:M83196; NID:g205537; PIDN:AAB48069.1; PID:g205538
A;Note: sequence extracted from NCBI backbone (NCBIN:111039, NCBIP:111040)
R;Cravchik, A.
submitted to the EMBL Data Library, June 1992
A;Reference number: S22108
A;Accession: S22108
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 73-364,'NRLRS',370,'QKN',374,'PSPKGL',381-751,'RSMMSQMNAQRR',764,'D',766,'L
'WLKRNMCPQPRQSP',851,'V',853,'NSL',855,'LPHRWLRTN',865,'W',867,'HSQLPDGGD',877,'Q',879,
A;Cross-references: EMBL:X66840
A;Experimental source: strain Sprague Dawley
C;Superfamily: microtubule-associated protein MAP1B
C;Keywords: microtubule binding; phosphoprotein

  Query Match             53.5%;  Score 46;  DB 2;  Length 2774;
  Best Local Similarity   87.5%;  Pred. No. 2.8e+02;
  Matches    7;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          7 PRPTPPRP 14
              |||:||||
Db       2543 PRPSPPRP 2550
```

```
RESULT 12
S54156
extensin-like protein - cowpea (fragment)
C;Species: Vigna unguiculata (cowpea)
C;Date: 08-Jul-1995 #sequence_revision 03-Aug-1995 #text_change 11-Jan-2000
C;Accession: S54156
R;Arsenijevic-Maksimovic, I.; Broughton, W.J.; Krause, A.
submitted to the EMBL Data Library, April 1995
A;Description: A class of root-hair specific extensins involved in rhizobium/legume inte
A;Reference number: S54155
A;Accession: S54156
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-242 <ARS>
A;Cross-references: EMBL:X86029; NID:g791147; PID:g791148
C;Superfamily: hydroxyproline-rich glycoprotein

  Query Match             52.3%;  Score 45;  DB 2;  Length 242;
  Best Local Similarity   70.0%;  Pred. No. 32;
  Matches    7;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          7 PRPTPPRPIY 16
              |||:|| | |
Db        127 PRPSPPPPYY 136


RESULT 13
T31742
hypothetical protein C05C8.4 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999
C;Accession: T31742
R;Sammons, L.; Wohldmann, P.
submitted to the EMBL Data Library, July 1997
A;Description: The sequence of C. elegans cosmid C05C8.
A;Reference number: Z21078
A;Accession: T31742
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-1106 <SAM>
A;Cross-references: EMBL:AF016430; PIDN:AAB65371.1; GSPDB:GN00023; CESP:C05C8.4
A;Experimental source: strain Bristol N2; clone C05C8
C;Genetics:
A;Gene: CESP:C05C8.4
A;Map position: 5
A;Introns: 25/3; 78/3; 117/1; 245/1; 591/1; 787/1; 1008/2

  Query Match             52.3%;  Score 45;  DB 2;  Length 1106;
  Best Local Similarity   77.8%;  Pred. No. 1.5e+02;
  Matches    7;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          7 PRPTPPRPI 15
              ||| ||||:
Db       1037 PRPMPPRPM 1045


RESULT 14
T31065
diaphanous protein homolog p140mDia - mouse
C;Species: Mus musculus (house mouse)
C;Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 22-Oct-1999
C;Accession: T31065
R;Watanabe, N.; Madaule, P.; Reid, T.; Ishizaki, T.; Watanabe, G.; Kakizuka, A.; Saito,
EMBO J. 16, 3044-3056, 1997
A;Title: P140mDia, a mammalian homolog of Drosophila diaphanous, is a target protein for
A;Reference number: Z20961; MUID:97357293; PMID:9214622
A;Accession: T31065
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-1255 <WAT>
A;Cross-references: EMBL:U96963; NID:g2114472; PID:g2114473; PIDN:AAC53280.1
A;Note: binds to GTP-bound form of Rho and binds to profilin

  Query Match             52.3%;  Score 45;  DB 2;  Length 1255;
  Best Local Similarity   46.7%;  Pred. No. 1.7e+02;
  Matches    7;  Conservative    2;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 DKGXXLPRPTPPRPI 15
              | |  :| | || |:
Db        580 DSGTVIPPPPPPPPL 594


RESULT 15
T13353
protein stn-B - fruit fly (Drosophila melanogaster)
C;Species: Drosophila melanogaster
C;Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 17-Nov-2000
C;Accession: T13353
R;Kelly, L.
submitted to the EMBL Data Library, May 1998
A;Reference number: Z17660
A;Accession: T13353
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-1262 <KEL>
A;Cross-references: EMBL:U54982; NID:g3138877; PID:g3138879; PIDN:AAC16666.1
C;Genetics:
A;Cross-references: FlyBase:FBgn0016976

  Query Match             52.3%;  Score 45;  DB 2;  Length 1262;
  Best Local Similarity   87.5%;  Pred. No. 1.7e+02;
  Matches    7;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          7 PRPTPPRP 14
              ||| ||||
Db        241 PRPAPPRP 248


Search completed: March 11, 2004, 16:54:06
Job time : 23 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: March 11, 2004, 16:53:29 ; Search time 11 Seconds
(without alignments)
85.206 Million cell updates/sec

Title: US-09-980-804-1
Perfect score: 86
Sequence: 1 DKGXXLPRPTPPRPIYXX 18

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	81	94.2	20	1	PYRR_PYRAP	P37362 pyrrhocoris
2	79	91.9	678	1	ABPP_RIPCL	Q27905 riptortus c
3	48	55.8	957	1	IF2_SYNEL	Q8dk04 synechococc
4	46	53.5	487	1	EBN2_EBV	P12978 epstein-bar
5	46	53.5	864	1	WS14_MOUSE	Q99mz3 mus musculu
6	46	53.5	902	1	NFC4_HUMAN	Q14934 homo sapien
7	46	53.5	2774	1	MAPA_RAT	P34926 rattus norv
8	46	53.5	3301	1	CLR3_MOUSE	Q91zi0 mus musculu
9	46	53.5	3312	1	CLR3_HUMAN	Q9nyq7 homo sapien
10	46	53.5	3313	1	CLR3_RAT	O88278 rattus norv
11	45	52.3	620	1	EXON_HSV2	P06489 herpes simp
12	45	52.3	1255	1	DIA1_MOUSE	O08808 mus musculu
13	45	52.3	1262	1	STNB_DROME	Q24212 drosophila
14	44	51.2	156	1	RS7_SYNY3	P74229 synechocyst
15	44	51.2	232	1	ACRL_HUMAN	P58840 homo sapien
16	44	51.2	250	1	EVG1_DROME	Q9vss7 drosophila
17	44	51.2	415	1	SYN1_CANFA	O62732 canis famil
18	44	51.2	421	1	ACRO_HUMAN	P10323 homo sapien
19	44	51.2	704	1	SYN1_RAT	P09951 rattus norv
20	44	51.2	705	1	SYN1_HUMAN	P17600 homo sapien
21	44	51.2	706	1	SYN1_BOVIN	P17599 bos taurus
22	44	51.2	742	1	PKWA_THECU	P49695 thermomonos
23	44	51.2	852	1	WS14_HUMAN	Q9np71 homo sapien
24	44	51.2	861	1	PQ58_CAEEL	P34552 caenorhabdi
25	44	51.2	899	1	SUHW_DROVI	Q08876 drosophila
26	43	50.0	351	1	ATH1_MOUSE	P48985 mus musculu
27	43	50.0	436	1	MOEA_ANASP	Q44243 anabaena sp
28	43	50.0	724	1	P85A_BOVIN	P23727 bos taurus
29	43	50.0	724	1	P85A_HUMAN	P27986 homo sapien
30	43	50.0	724	1	P85A_MOUSE	P26450 mus musculu
31	43	50.0	724	1	P85A_RAT	Q63787 rattus norv
32	43	50.0	786	1	CT32_HUMAN	Q9nq75 homo sapien
33	43	50.0	925	1	UVRA_ZYMMO	O31151 zymomonas m
34	43	50.0	1003	1	P5PA_MOUSE	P59644 mus musculu
35	43	50.0	1131	1	RPOB_ANASP	P22703 anabaena sp
36	43	50.0	1183	1	YNK6_YEAST	P50942 saccharomyc
37	43	50.0	3767	1	MUA3_CAEEL	P34576 caenorhabdi
38	42	48.8	179	1	LEB1_BOMMO	P54684 bombyx mori
39	42	48.8	197	1	HPAP_RALSO	P35651 ralstonia s
40	42	48.8	385	1	RURE_PSEOL	P17052 pseudomonas
41	42	48.8	389	1	SH2A_HUMAN	Q9np31 homo sapien
42	42	48.8	487	1	HEP_DROME	Q23977 drosophila
43	42	48.8	491	1	YO84_MYCTU	O53209 mycobacteri
44	42	48.8	577	1	MTG8_MOUSE	Q61909 mus musculu
45	42	48.8	604	1	MTG8_HUMAN	Q06455 homo sapien

ALIGNMENTS

RESULT 1
PYRR_PYRAP
ID PYRR_PYRAP STANDARD; PRT; 20 AA.
AC P37362; P80307;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Pyrrhocoricin.
OS Pyrrhocoris apterus (Sap sucking bug).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Paraneoptera; Hemiptera; Euhemiptera; Heteroptera;
OC Panheteroptera; Pentatomomorpha; Pyrrhocoroidea; Pyrrhocoridae;
OC Pyrrhocoris.
OX NCBI_TaxID=37000;
RN [1]
RP SEQUENCE.
RC TISSUE=Hemolymph;
RX MEDLINE=94271176; PubMed=8002963;
RA Cociancich S., Dupont A., Hegy G., Lanot R., Holder F., Hetru C.,
RA Hoffmann J.A., Bulet P.;
RT "Novel inducible antibacterial peptides from a hemipteran insect, the
RT sap-sucking bug Pyrrhocoris apterus.";
RL Biochem. J. 300:567-575(1994).
RN [2]
RP CARBOHYDRATE-LINKAGE SITE THR-11.
RX MEDLINE=99177428; PubMed=10076062;
RA Hoffmann R., Bulet P., Urge L., Otvoes L. Jr.;
RT "Range of activity and metabolic stability of synthetic antibacterial
RT glycopeptides from insects.";
RL Biochim. Biophys. Acta 1426:459-467(1999).
CC -!- FUNCTION: Antibacterial peptide. Affects Gram-negative bacteria
CC E.coli 1106, P.aeruginosa, E.coli D22 and E.cloacae and
CC Gram-positive bacteria M.luteus and B.subtilis.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- PTM: O-LINKED GLYCAN CONSISTS OF A GAL-GALNAC DISACCHARIDE, O-
CC GLYCOSYLATION IS ESSENTIAL FOR FULL BIOLOGICAL ACTIVITY.
CC -!- SIMILARITY: TO DROSOPHILA DROSOCIN.
DR PIR; S44465; S44465.
KW Antibiotic; Glycoprotein; Insect immunity; Hemolymph.
FT CARBOHYD 11 11 O-LINKED (GALNAC. . .).
SQ SEQUENCE 20 AA; 2341 MW; F4320EC2FF29462C CRC64;

Query Match 94.2%; Score 81; DB 1; Length 20;
Best Local Similarity 87.5%; Pred. No. 2.6e-05;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy          1 DKGXXLPRPTPPRPIY 16
              ||| |||||||||||
Db          2 DKGSYLPRPTPPRPIY 17
```

RESULT 2
ABPP_RIPCL
ID ABPP_RIPCL STANDARD; PRT; 678 AA.
AC Q27905;
DT 01-NOV-1997 (Rel. 35, Created)

```
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Probable antibacterial peptide polyprotein precursor.
OS   Riptortus clavatus (Bean bug).
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Paraneoptera; Hemiptera; Euhemiptera; Heteroptera;
OC   Panheteroptera; Pentatomomorpha; Coreoidea; Alydidae; Riptortus.
OX   NCBI_TaxID=41704;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Fat body;
RX   MEDLINE=96261233; PubMed=8688805;
RA   Miura K., Ueno S., Kamiya K., Kobayashi J., Matsuoka H., Ando K.,
RA   Chinzei Y.;
RT   "Cloning of mRNA sequences for two antibacterial peptides in a
RT   hemipteran insect, Riptortus clavatus.";
RL   Zool. Sci. 13:111-117(1996).
CC   -!- FUNCTION: Has antibacterial activity in vitro.
CC   -!- SUBCELLULAR LOCATION: Secreted (Potential).
CC   -!- SIMILARITY: TO PYRRHOCORICIN, DROSOCIN AND APIDAECIN.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; D49415; BAA08395.1; -.
KW   Polyprotein; Insect immunity; Antibiotic; Repeat; Glycoprotein.
FT   DOMAIN        1    678       14 X APPROXIMATE TANDEM REPEATS.
FT   REPEAT        1     67       1-1.
FT   REPEAT       68    114       1-2.
FT   REPEAT      115    161       1-3.
FT   REPEAT      162    208       1-4.
FT   REPEAT      209    255       1-5.
FT   REPEAT      256    302       1-6.
FT   REPEAT      303    349       1-7.
FT   REPEAT      350    396       1-8.
FT   REPEAT      397    443       1-9.
FT   REPEAT      444    490       1-10.
FT   REPEAT      491    537       1-11.
FT   REPEAT      538    584       1-12.
FT   REPEAT      585    631       1-13.
FT   REPEAT      632    678       1-14.
FT   CARBOHYD     32     32       O-LINKED (GALNAC. . .) (POTENTIAL).
FT   CARBOHYD     83     83       O-LINKED (GALNAC. . .) (POTENTIAL).
FT   CARBOHYD    130    130       O-LINKED (GALNAC. . .) (POTENTIAL).
FT   CARBOHYD    177    177       O-LINKED (GALNAC. . .) (POTENTIAL).
FT   CARBOHYD    224    224       O-LINKED (GALNAC. . .) (POTENTIAL).
FT   CARBOHYD    271    271       O-LINKED (GALNAC. . .) (POTENTIAL).
FT   CARBOHYD    318    318       O-LINKED (GALNAC. . .) (POTENTIAL).
FT   CARBOHYD    365    365       O-LINKED (GALNAC. . .) (POTENTIAL).
FT   CARBOHYD    412    412       O-LINKED (GALNAC. . .) (POTENTIAL).
FT   CARBOHYD    459    459       O-LINKED (GALNAC. . .) (POTENTIAL).
FT   CARBOHYD    506    506       O-LINKED (GALNAC. . .) (POTENTIAL).
FT   CARBOHYD    553    553       O-LINKED (GALNAC. . .) (POTENTIAL).
FT   CARBOHYD    600    600       O-LINKED (GALNAC. . .) (POTENTIAL).
FT   CARBOHYD    647    647       O-LINKED (GALNAC. . .) (POTENTIAL).
SQ   SEQUENCE   678 AA;  76367 MW;  2939BA68692D2444 CRC64;

  Query Match             91.9%;  Score 79;  DB 1;  Length 678;
  Best Local Similarity   81.2%;  Pred. No. 0.0016;
  Matches   13;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 DKGXXLPRPTPPRPIY 16
              |||  |||||||||:|
Db         74 DKGGYLPRPTPPRPVY 89


RESULT 3
IF2_SYNEL
```

```
ID   IF2_SYNEL      STANDARD;      PRT;   957 AA.
AC   Q8DK04;
DT   15-MAR-2004 (Rel. 43, Created)
DT   15-MAR-2004 (Rel. 43, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   Translation initiation factor IF-2.
GN   INFB OR TLR1066.
OS   Synechococcus elongatus (Thermosynechococcus elongatus).
OC   Bacteria; Cyanobacteria; Chroococcales; Synechococcus.
OX   NCBI_TaxID=32046;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BP-1;
RX   MEDLINE=22225144; PubMed=12240834;
RA   Nakamura Y., Kaneko T., Sato S., Ikeuchi M., Katoh H., Sasamoto S.,
RA   Watanabe A., Iriguchi M., Kawashima K., Kimura T., Kishida Y.,
RA   Kiyokawa C., Kohara M., Matsumoto M., Matsuno A., Nakazaki N.,
RA   Shimpo S., Sugimoto M., Takeuchi C., Yamada M., Tabata S.;
RT   "Complete genome structure of the thermophilic cyanobacterium
RT   Thermosynechococcus elongatus BP-1.";
RL   DNA Res. 9:123-130(2002).
CC   -!- FUNCTION: One of the essential components for the initiation of
CC       protein synthesis. Protects formylmethionyl-tRNA from spontaneous
CC       hydrolysis and promotes its binding to the 30S ribosomal subunits.
CC       Also involved in the hydrolysis of GTP during the formation of the
CC       70S ribosomal complex (By similarity).
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC   -!- SIMILARITY: Belongs to the IF-2 family.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AP005372; BAC08619.1; -.
DR   HAMAP; MF_00100; -; 1.
DR   InterPro; IPR004161; EFTU_D2.
DR   InterPro; IPR000795; EF_GTPbind.
DR   InterPro; IPR000178; IF2.
DR   InterPro; IPR006847; IF2_N.
DR   InterPro; IPR002965; P_rich_extensn.
DR   InterPro; IPR005225; Small_GTP.
DR   InterPro; IPR009000; Translat_factor.
DR   Pfam; PF00009; GTP_EFTU; 1.
DR   Pfam; PF03144; GTP_EFTU_D2; 2.
DR   Pfam; PF04760; IF2_N; 2.
DR   PRINTS; PR00315; ELONGATNFCT.
DR   PRINTS; PR01217; PRICHEXTENSN.
DR   ProDom; PD186100; IF2; 1.
DR   TIGRFAMs; TIGR00487; IF-2; 1.
DR   TIGRFAMs; TIGR00231; small_GTP; 1.
DR   PROSITE; PS01176; IF2; 1.
KW   Initiation factor; Protein biosynthesis; GTP-binding;
KW   Complete proteome.
FT   DOMAIN      447    599       G-DOMAIN.
FT   NP_BIND     453    460       GTP (BY SIMILARITY).
FT   NP_BIND     503    507       GTP (BY SIMILARITY).
FT   NP_BIND     557    560       GTP (BY SIMILARITY).
SQ   SEQUENCE   957 AA;  104247 MW;  13E9E041ADBC1280 CRC64;

  Query Match             55.8%;  Score 48;  DB 1;  Length 957;
  Best Local Similarity   77.8%;  Pred. No. 31;
  Matches    7;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          7 PRPTPPRPI 15
              | ||||||:
Db        144 PEPTPPRPV 152


RESULT 4
```

```
EBN2_EBV
ID   EBN2_EBV       STANDARD;      PRT;   487 AA.
AC   P12978;
DT   01-OCT-1989 (Rel. 12, Created)
DT   01-OCT-1989 (Rel. 12, Last sequence update)
DT   01-NOV-1995 (Rel. 32, Last annotation update)
DE   EBNA-2 nuclear protein.
GN   BYRF1.
OS   Epstein-barr virus (strain B95-8) (Human herpesvirus 4).
OC   Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC   Gammaherpesvirinae; Lymphocryptovirus.
OX   NCBI_TaxID=10377;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=84270667; PubMed=6087149;
RA   Baer R., Bankier A.T., Biggin M.D., Deininger P.L., Farrell P.J.,
RA   Gibson T.J., Hatfull G., Hudson G.S., Satchwell S.C., Seguin C.,
RA   Tuffnell P.S., Barrell B.G.;
RT   "DNA sequence and expression of the B95-8 Epstein-Barr virus genome.";
RL   Nature 310:207-211(1984).
RN   [2]
RP   SUBCELLULAR LOCATION, AND PHOSPHORYLATION.
RX   MEDLINE=90266473; PubMed=2161150;
RA   Petti L., Sample C., Kieff E.;
RT   "Subnuclear localization and phosphorylation of Epstein-Barr virus
RT   latent infection nuclear proteins.";
RL   Virology 176:563-574(1990).
RN   [3]
RP   DOMAINS.
RX   MEDLINE=91202599; PubMed=1850028;
RA   Cohen J.I., Wang F., Kieff E.;
RT   "Epstein-Barr virus nuclear protein 2 mutations define essential
RT   domains for transformation and transactivation.";
RL   J. Virol. 65:2545-2554(1991).
CC   -!- FUNCTION: INVOLVED IN LATENT CYCLE. TRANSACTIVATES THE EXPRESSION
CC       OF LMP-1.
CC   -!- SUBCELLULAR LOCATION: Nuclear. Associated with the nuclear matrix.
CC   -!- PTM: Phosphorylated.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; V01555; CAA24877.1; ALT_INIT.
DR   PIR; S42442; S42442.
DR   TRANSFAC; T01618; -.
KW   Transcription regulation; Activator; Nuclear protein; DNA-binding;
KW   Phosphorylation; Repeat.
FT   DOMAIN       59    100       POLY-PRO.
FT   DOMAIN      345    356       6 X 2 AA TANDEM REPEATS OF R-G.
SQ   SEQUENCE   487 AA;  52544 MW;  DEF40D7F8ED61D1A CRC64;

  Query Match             53.5%;  Score 46;  DB 1;  Length 487;
  Best Local Similarity   77.8%;  Pred. No. 29;
  Matches    7;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          7 PRPTPPRPI 15
              |||||| |:
Db        198 PRPTPPTPL 206


RESULT 5
WS14_MOUSE
ID   WS14_MOUSE     STANDARD;      PRT;   864 AA.
AC   Q99MZ3; Q99MY9; Q99MZ0; Q99MZ1; Q99MZ2; Q9JLM5;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Williams-Beuren syndrome chromosome region 14 protein homolog (Mlx
DE   interactor).
GN   WBSCR14 OR MIO.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A., CHARACTERIZATION, AND ALTERNATIVE SPLICING.
RX   MEDLINE=21153101; PubMed=11230181;
RA   Cairo S., Merla G., Urbinati F., Ballabio A., Reymond A.;
RT   "WBSCR14, a gene mapping to the Williams-Beuren syndrome deleted
RT   region, is a new member of the Mlx transcription factor network.";
RL   Hum. Mol. Genet. 10:617-627(2001).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20241700; PubMed=10780788;
RA   de Luis O., Valero M.C., Perez Jurado L.A.;
RT   "WBSCR14, a putative transcription factor gene deleted in Williams-
RT   Beuren syndrome: complete characterisation of the human gene and the
RT   mouse ortholog.";
RL   Eur. J. Hum. Genet. 8:215-222(2000).
CC   -!- FUNCTION: Transcriptional repressor. Binds to the canonical and
CC       non-canonical E box sequences 5'-CACGTG-3'.
CC   -!- SUBUNIT: Binds DNA as a heterodimer with TCFL4/MLX.
CC   -!- SUBCELLULAR LOCATION: Nuclear (By similarity).
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=5;
CC       Name=1; Synonyms=Zeta;
CC         IsoId=Q99MZ3-1; Sequence=Displayed;
CC       Name=2; Synonyms=Theta;
CC         IsoId=Q99MZ3-2; Sequence=VSP_002174;
CC       Name=3; Synonyms=Iota;
CC         IsoId=Q99MZ3-3; Sequence=VSP_002177, VSP_002178;
CC       Name=4; Synonyms=Kappa;
CC         IsoId=Q99MZ3-4; Sequence=VSP_002179, VSP_002180;
CC       Name=5; Synonyms=Eta;
CC         IsoId=Q99MZ3-5; Sequence=VSP_002175, VSP_002176;
CC   -!- TISSUE SPECIFICITY: Expressed in the ventricular and intermediate
CC       zones of the developing spinal cord of E12.5 embryos. In later
CC       embryos expressed in a variety of tissues.
CC   -!- SIMILARITY: Contains 1 basic helix-loop-helix (bHLH) domain.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF245475; AAK20940.1; -.
DR   EMBL; AF245476; AAK20941.1; -.
DR   EMBL; AF245477; AAK20942.1; -.
DR   EMBL; AF245478; AAK20943.1; -.
DR   EMBL; AF245479; AAK20944.1; -.
DR   EMBL; AF156604; AAF68175.1; -.
DR   HSSP; P25912; 1HLO.
DR   TRANSFAC; T05122; -.
DR   MGD; MGI:1927999; Wbscr14.
DR   GO; GO:0005667; C:transcription factor complex; IDA.
DR   GO; GO:0016564; F:transcriptional repressor activity; IDA.
DR   GO; GO:0000122; P:negative regulation of transcription from P. . .; IDA.
DR   InterPro; IPR001092; HLH_basic.
DR   Pfam; PF00010; HLH; 1.
DR   SMART; SM00353; HLH; 1.
DR   PROSITE; PS50888; HLH; 1.
KW   Transcription regulation; Repressor; Nuclear protein; DNA-binding;
KW   Alternative splicing.
FT   DOMAIN      345    350       POLY-SER.
FT   DNA_BIND    660    674       BASIC DOMAIN.
FT   DOMAIN      700    714       HELIX-LOOP-HELIX MOTIF.
FT   DOMAIN      715    736       LEUCINE-ZIPPER.
FT   VARSPLIC     58     79       Missing (in isoform 2).
```

```
FT                                     /FTId=VSP_002174.
FT   VARSPLIC    545    556        AKPEQALEPPTM -> VVLIVLPVPSQA (in isoform
FT                                 5).
FT                                     /FTId=VSP_002175.
FT   VARSPLIC    557    864        Missing (in isoform 5).
FT                                     /FTId=VSP_002176.
FT   VARSPLIC    699    744        VSKATTLQKTAEYILMLQQERAAMQEEAQQLRDEIEELNAA
FT                                 INLCQ -> GLPTQRPTLVALAGEQSNHASEDSGVHPDAAA
FT                                 GTGSYAGGGAAAAG (in isoform 3).
FT                                     /FTId=VSP_002177.
FT   VARSPLIC    745    864        Missing (in isoform 3).
FT                                     /FTId=VSP_002178.
FT   VARSPLIC    699    714        VSKATTLQKTAEYILM -> LPGLANTEAHIGGARR (in
FT                                 isoform 4).
FT                                     /FTId=VSP_002179.
FT   VARSPLIC    715    864        Missing (in isoform 4).
FT                                     /FTId=VSP_002180.
FT   CONFLICT     67     67        D -> Y (IN REF. 2).
FT   CONFLICT    107    107        K -> N (IN REF. 2).
FT   CONFLICT    128    128        R -> I (IN REF. 2).
FT   CONFLICT    138    139        RK -> TR (IN REF. 2).
FT   CONFLICT    155    155        D -> H (IN REF. 2).
FT   CONFLICT    175    175        E -> D (IN REF. 2).
FT   CONFLICT    183    183        K -> V (IN REF. 2).
FT   CONFLICT    727    728        QQ -> HE (IN REF. 2).
SQ   SEQUENCE   864 AA;  94874 MW;  7E6AFFB04C71B327 CRC64;

  Query Match             53.5%;  Score 46;  DB 1;  Length 864;
  Best Local Similarity   77.8%;  Pred. No. 51;
  Matches    7;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          6 LPRPTPPRP 14
              :| ||||||
Db        584 IPAPTPPRP 592


RESULT 6
NFC4_HUMAN
ID   NFC4_HUMAN     STANDARD;      PRT;   902 AA.
AC   Q14934;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   Nuclear factor of activated T-cells, cytoplasmic 4 (T cell
DE   transcription factor NFAT3) (NF-ATc4) (NF-AT3).
GN   NFATC4 OR NFAT3.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=T-cell;
RX   MEDLINE=95269130; PubMed=7749981;
RA   Hoey T., Sun Y.-L., Williamson K., Xu X.;
RT   "Isolation of two new members of the NF-AT gene family and functional
RT   characterization of the NF-AT proteins.";
RL   Immunity 2:461-472(1995).
RN   [2]
RP   REVIEW.
RX   MEDLINE=99189746; PubMed=10089876;
RA   Crabtree G.R.;
RT   "Generic signals and specific outcomes: signaling through Ca2+,
RT   calcineurin, and NF-AT.";
RL   Cell 96:611-614(1999).
CC   -!- FUNCTION: Plays a role in the inducible expression of cytokine
CC       genes in T cells, especially in the induction of the IL-2 and IL-
CC       4 (By similarity).
CC   -!- SUBUNIT: Member of the multicomponent NFATC transcription complex
CC       that consists of at least two components, a pre-existing
CC       cytoplasmic component NFATC2 and an inducible nuclear component
CC       NFATC1. Other members such as NFATC4, NFATC3 or members of the
CC       activating protein-1 family, MAF, GATA4 and Cbp/p300 can also bind
CC       the complex. NFATC proteins bind to DNA as monomers.
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic for the phosphorylated form and
CC       nuclear after activation that is controlled by calcineurin-
CC       mediated dephosphorylation. Rapid nuclear exit of NFATC is thought
CC       to be one mechanism by which cells distinguish between sustained
CC       and transient calcium signals. The subcellular localization of
CC       NFATC play a key role in the gene transcription.
CC   -!- TISSUE SPECIFICITY: Highly expressed in placenta, lung, kidney,
CC       testis and ovary. Weakly expressed in spleen and thymus. Not
CC       expressed in peripheral blood lymphocytes.
CC   -!- DOMAIN: Rel Similarity Domain (RSD) allows DNA-binding and
CC       cooperative interactions with AP1 factors (By similarity).
CC   -!- PTM: Phosphorylated by NFATC-kinase; dephosphorylated by
CC       calcineurin (By similarity).
CC   -!- SIMILARITY: Belongs to the Rel/Dorsal family.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; L41066; AAA79175.1; -.
DR   TRANSFAC; T02462; -.
DR   Genew; HGNC:7778; NFATC4.
DR   MIM; 602699; -.
DR   GO; GO:0003713; F:transcription co-activator activity; TAS.
DR   GO; GO:0006954; P:inflammatory response; TAS.
DR   GO; GO:0006366; P:transcription from Pol II promoter; TAS.
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR002909; IPT_TIG.
DR   InterPro; IPR000451; NF_Rel_dor.
DR   InterPro; IPR008366; NFAT.
DR   InterPro; IPR008967; P53-like.
DR   Pfam; PF00554; RHD; 1.
DR   Pfam; PF01833; TIG; 1.
DR   PRINTS; PR01789; NUCFACTORATC.
DR   SMART; SM00429; IPT; 1.
DR   PROSITE; PS01204; REL_1; FALSE_NEG.
DR   PROSITE; PS50254; REL_2; 1.
KW   Transcription regulation; Activator; Nuclear protein; DNA-binding;
KW   Repeat; Phosphorylation.
FT   DOMAIN       62     69       POLY-PRO.
FT   DOMAIN      114    119       CALCINEURIN-BINDING.
FT   DOMAIN      213    293       2 APPROXIMATE SP REPEATS.
FT   REPEAT      213    229       SP 1.
FT   REPEAT      277    293       SP 2 (APPROXIMATE).
FT   DOMAIN      297    304       POLY-PRO.
FT   DOMAIN      268    270       NUCLEAR LOCALIZATION SIGNAL.
FT   DOMAIN      430    437       DNA-BINDING.
FT   DOMAIN      672    674       NUCLEAR LOCALIZATION SIGNAL.
SQ   SEQUENCE   902 AA;  95472 MW;  E59F15F7647A47C6 CRC64;

  Query Match             53.5%;  Score 46;  DB 1;  Length 902;
  Best Local Similarity   77.8%;  Pred. No. 53;
  Matches    7;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          6 LPRPTPPRP 14
              :||| ||||
Db         61 IPRPPPPRP 69


RESULT 7
MAPA_RAT
ID   MAPA_RAT       STANDARD;      PRT;  2774 AA.
AC   P34926;
DT   01-FEB-1994 (Rel. 28, Created)
DT   01-FEB-1994 (Rel. 28, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Microtubule-associated protein 1A (MAP 1A) [Contains: MAP1 light chain
DE   LC2].
```

```
GN   MAP1A.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Brain;
RX   MEDLINE=92355629; PubMed=1379599;
RA   Langkopf A., Hammarback J.A., Mueller R., Vallee R.B., Garner C.C.;
RT   "Microtubule-associated proteins 1A and LC2. Two proteins encoded in
RT   one messenger RNA.";
RL   J. Biol. Chem. 267:16561-16566(1992).
CC   -!- FUNCTION: Structural protein involved in the filamentous cross-
CC       bridging between microtubules and other skeletal elements.
CC   -!- SUBUNIT: 3 different light chains, LC1, LC2 and LC3, can associate
CC       with MAP1A and MAP1B proteins.
CC   -!- TISSUE SPECIFICITY: BRAIN, HEART AND MUSCLE.
CC   -!- DEVELOPMENTAL STAGE: EXPRESSED LATE DURING NEURONAL DEVELOPMENT
CC       APPEARING WHEN AXONS AND DENDRITES BEGIN TO SOLIDIFY AND STABILIZE
CC       THEIR MORPHOLOGY.
CC   -!- DOMAIN: THe basic region containing the repeats may be responsible
CC       for the binding of MAP1A to microtubules.
CC   -!- PTM: Various serine residues may be phosphorylated by cAMP kinase.
CC   -!- PTM: LC2 IS COEXPRESSED WITH MAP1A. IT IS A POLYPEPTIDE GENERATED
CC       FROM MAP1A BY PROTEOLYTIC PROCESSING. IT IS FREE TO ASSOCIATE WITH
CC       BOTH MAP1A AND MAP1B.
CC   -!- SIMILARITY: TO MAP1B.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M83196; AAB48069.1; -.
DR   PIR; A43359; A43359.
KW   Microtubule; Repeat; Phosphorylation.
FT   CHAIN     ?2465   2774       MAP1 LIGHT CHAIN LC2.
FT   DOMAIN       309    496       LYS-RICH (BASIC).
FT   DOMAIN       336    541       11 X 3 AA REPEATS OF K-K-[DE].
FT   REPEAT       336    338       1.
FT   REPEAT       415    417       2.
FT   REPEAT       420    422       3.
FT   REPEAT       424    426       4.
FT   REPEAT       427    429       5.
FT   REPEAT       431    433       6.
FT   REPEAT       436    438       7.
FT   REPEAT       440    442       8.
FT   REPEAT       444    446       9.
FT   REPEAT       449    451       10.
FT   REPEAT       539    541       11.
SQ   SEQUENCE   2774 AA;  299526 MW;  3DEF74427BA9D7D7 CRC64;

  Query Match             53.5%;  Score 46;  DB 1;  Length 2774;
  Best Local Similarity   87.5%;  Pred. No. 1.6e+02;
  Matches    7;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          7 PRPTPPRP 14
              |||:||||
Db       2543 PRPSPPRP 2550


RESULT 8
CLR3_MOUSE
ID   CLR3_MOUSE     STANDARD;      PRT;  3301 AA.
AC   Q91ZI0; Q9ESD0;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Cadherin EGF LAG seven-pass G-type receptor 3 precursor.
GN   CELSR3.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A., AND DEVELOPMENTAL STAGES.
RC   STRAIN=C57BL/6;
RX   MEDLINE=21839555; PubMed=11850187;
RA   Tissir F., De-Backer O., Goffinet A.M., Lambert de Rouvroit C.A.;
RT   "Developmental expression profiles of Celsr (Flamingo) genes in the
RT   mouse.";
RL   Mech. Dev. 112:157-160(2002).
RN   [2]
RP   SEQUENCE OF 2574-3046 FROM N.A., AND DEVELOPMENTAL STAGE.
RX   MEDLINE=21534880; PubMed=11677057;
RA   Formstone C.J., Little P.F.R.;
RT   "The flamingo-related mouse Celsr family (Celsr1-3) genes exhibit
RT   distinct patterns of expression during embryonic development.";
RL   Mech. Dev. 109:91-94(2001).
RN   [3]
RP   TISSUE SPECIFICITY.
RX   MEDLINE=20253755; PubMed=10790539;
RA   Formstone C.J., Barclay J., Rees M., Little P.F.R.;
RT   "Chromosomal localization of Celsr2 and Celsr3 in the mouse; Celsr3 is
RT   a candidate for the tippy (tip) lethal mutant on chromosome 9.";
RL   Mamm. Genome 11:392-394(2000).
CC   -!- FUNCTION: Receptor that may have an important role in cell/cell
CC       signaling during nervous system formation.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- TISSUE SPECIFICITY: Expressed in the CNS and in the eye.
CC   -!- DEVELOPMENTAL STAGE: Predominantly expressed in the CNS, the
CC       emerging dorsal root ganglia and cranial ganglia. In the CNS,
CC       expression is uniform along the rostrocaudal axis. No expression
CC       is detected until somite stages. Between E10 and E12, expression
CC       is strong in the marginal zone (MZ), and lower in the ventricular
CC       zone (VZ). At E15, expression is restricted to the brain and
CC       olfactory epithelium. In the brain, it is low in VZ but strong in
CC       external fields, particularly those with ongoing migration, such
CC       as the telencephalic cortical plate, the olfactory bulb, the
CC       cerebrellum and the tectum. In the newborn and postnatal stages,
CC       expression is high in differenciated neuronal fields.
CC   -!- SIMILARITY: Belongs to family 2 of G-protein coupled receptors.
CC   -!- SIMILARITY: Contains 9 cadherin domains.
CC   -!- SIMILARITY: Contains 7 EGF-like domains.
CC   -!- SIMILARITY: Contains 2 laminin G-like domains.
CC   -!- SIMILARITY: Contains 1 laminin EGF-like domain.
CC   -!- SIMILARITY: Contains 1 GPS domain.
CC   -!- CAUTION: Ref.2 sequence differs from that shown due to frameshifts
CC       in positions 2575 and 2578.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF427498; AAL25099.1; -.
DR   EMBL; AF188752; AAG17057.1; ALT_FRAME.
DR   MGD; MGI:1858236; Celsr3.
DR   InterPro; IPR000152; Asx_hydroxyl_S.
DR   InterPro; IPR002126; Cadherin.
DR   InterPro; IPR008985; ConA_like_lec_gl.
DR   InterPro; IPR000742; EGF_2.
DR   InterPro; IPR006209; EGF_like.
DR   InterPro; IPR000832; GPCR_secretin.
DR   InterPro; IPR001879; hormn_receptor.
DR   InterPro; IPR006210; IEGF.
DR   InterPro; IPR002049; Laminin_EGF.
DR   InterPro; IPR001791; Laminin_G.
DR   InterPro; IPR000203; PKD_cys_rich.
```

```
DR   Pfam; PF00002; 7tm_2; 1.
DR   Pfam; PF00028; cadherin; 9.
DR   Pfam; PF00008; EGF; 5.
DR   Pfam; PF01825; GPS; 1.
DR   Pfam; PF02793; HRM; 1.
DR   Pfam; PF00053; laminin_EGF; 1.
DR   Pfam; PF00054; laminin_G; 2.
DR   PRINTS; PR00205; CADHERIN.
DR   PRINTS; PR00249; GPCRSECRETIN.
DR   SMART; SM00112; CA; 9.
DR   SMART; SM00181; EGF; 6.
DR   SMART; SM00303; GPS; 1.
DR   SMART; SM00008; HormR; 1.
DR   SMART; SM00282; LamG; 2.
DR   PROSITE; PS00010; ASX_HYDROXYL; 1.
DR   PROSITE; PS00232; CADHERIN_1; 7.
DR   PROSITE; PS50268; CADHERIN_2; 8.
DR   PROSITE; PS00022; EGF_1; 5.
DR   PROSITE; PS01186; EGF_2; 4.
DR   PROSITE; PS50026; EGF_3; 6.
DR   PROSITE; PS00649; G_PROTEIN_RECEP_F2_1; FALSE_NEG.
DR   PROSITE; PS00650; G_PROTEIN_RECEP_F2_2; FALSE_NEG.
DR   PROSITE; PS50227; G_PROTEIN_RECEP_F2_3; 1.
DR   PROSITE; PS50261; G_PROTEIN_RECEP_F2_4; 1.
DR   PROSITE; PS50221; GPS; 1.
DR   PROSITE; PS50025; LAM_G_DOMAIN; 2.
DR   PROSITE; PS01248; LAMININ_TYPE_EGF; 1.
KW   G-protein coupled receptor; Transmembrane; Glycoprotein;
KW   EGF-like domain; Calcium-binding; Laminin EGF-like domain; Repeat;
KW   Developmental protein; Hydroxylation; Signal.
FT   SIGNAL        1     31       POTENTIAL.
FT   CHAIN        32   3301       CADHERIN EGF LAG SEVEN-PASS G-TYPE
FT                                RECEPTOR 3.
FT   DOMAIN       32   2531       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   2532   2552       1 (POTENTIAL).
FT   DOMAIN     2553   2563       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM   2564   2584       2 (POTENTIAL).
FT   DOMAIN     2585   2592       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   2593   2613       3 (POTENTIAL).
FT   DOMAIN     2614   2634       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM   2635   2655       4 (POTENTIAL).
FT   DOMAIN     2656   2673       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   2674   2694       5 (POTENTIAL).
FT   DOMAIN     2695   2716       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM   2717   2737       6 (POTENTIAL).
FT   DOMAIN     2738   2744       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   2745   2765       7 (POTENTIAL).
FT   DOMAIN     2766   3301       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN      317    424       CADHERIN 1.
FT   DOMAIN      425    536       CADHERIN 2.
FT   DOMAIN      537    642       CADHERIN 3.
FT   DOMAIN      643    747       CADHERIN 4.
FT   DOMAIN      748    849       CADHERIN 5.
FT   DOMAIN      850    952       CADHERIN 6.
FT   DOMAIN      953   1058       CADHERIN 7.
FT   DOMAIN     1059   1160       CADHERIN 8.
FT   DOMAIN     1161   1257       CADHERIN 9.
FT   DOMAIN     1366   1424       EGF-LIKE 1, CALCIUM-BINDING.
FT   DOMAIN     1426   1460       EGF-LIKE 2, CALCIUM-BINDING.
FT   DOMAIN     1464   1503       EGF-LIKE 3, CALCIUM-BINDING.
FT   DOMAIN     1504   1708       LAMININ G-LIKE 1.
FT   DOMAIN     1711   1747       EGF-LIKE 4, CALCIUM-BINDING.
FT   DOMAIN     1751   1933       LAMININ G-LIKE 2.
FT   DOMAIN     1935   1971       EGF-LIKE 5, CALCIUM-BINDING.
FT   DOMAIN     1972   2002       EGF-LIKE 6, CALCIUM-BINDING.
FT   DOMAIN     2003   2042       EGF-LIKE 7, CALCIUM-BINDING.
FT   DOMAIN     2044   2079       EGF-LIKE 8, CALCIUM-BINDING.
FT   DOMAIN     2085   2118       LAMININ EGF-LIKE.
FT   DOMAIN     2468   2520       GPS.
FT   DOMAIN     2720   2724       POLY-LEU.
FT   DISULFID   1370   1381       BY SIMILARITY.
FT   DISULFID   1375   1412       BY SIMILARITY.
FT   DISULFID   1414   1423       BY SIMILARITY.
FT   DISULFID   1430   1441       BY SIMILARITY.
FT   DISULFID   1435   1450       BY SIMILARITY.
FT   DISULFID   1452   1459       BY SIMILARITY.
FT   DISULFID   1468   1479       BY SIMILARITY.
FT   DISULFID   1473   1489       BY SIMILARITY.
FT   DISULFID   1491   1502       BY SIMILARITY.
FT   DISULFID   1715   1726       BY SIMILARITY.
FT   DISULFID   1720   1735       BY SIMILARITY.
FT   DISULFID   1737   1746       BY SIMILARITY.
FT   DISULFID   1939   1950       BY SIMILARITY.
FT   DISULFID   1944   1959       BY SIMILARITY.
FT   DISULFID   1961   1970       BY SIMILARITY.
FT   DISULFID   1974   1985       BY SIMILARITY.
FT   DISULFID   1979   1997       BY SIMILARITY.
FT   DISULFID   1999   2008       BY SIMILARITY.
FT   DISULFID   2016   2029       BY SIMILARITY.
FT   DISULFID   2031   2041       BY SIMILARITY.
FT   DISULFID   2048   2063       BY SIMILARITY.
FT   DISULFID   2050   2066       BY SIMILARITY.
FT   DISULFID   2068   2078       BY SIMILARITY.
FT   MOD_RES    1952   1952       HYDROXYLATION (POTENTIAL).
FT   CARBOHYD    623    623       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    838    838       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1173   1173       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1213   1213       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1308   1308       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1318   1318       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1638   1638       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1702   1702       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1759   1759       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   2042   2042       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   2166   2166       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   2185   2185       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   2375   2375       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   2465   2465       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   2497   2497       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CONFLICT   2713   2713       L -> LR (IN REF. 2).
FT   CONFLICT   3024   3024       R -> P (IN REF. 2).
SQ   SEQUENCE   3301 AA;  358455 MW;  A6B18F2DF7F4DEB6 CRC64;

  Query Match           53.5%;  Score 46;  DB 1;  Length 3301;
  Best Local Similarity   61.5%;  Pred. No. 1.9e+02;
  Matches    8;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 DKGXXLPRPTPPR 13
              |:|  |||  |||
Db       3109 DRGSTLPRRQPPR 3121


RESULT 9
CLR3_HUMAN
ID   CLR3_HUMAN     STANDARD;      PRT;  3312 AA.
AC   Q9NYQ7; O75092;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Cadherin EGF LAG seven-pass G-type receptor 3 precursor (Flamingo
DE   homolog 1) (hFmi1) (Multiple epidermal growth factor-like domains 2)
DE   (Epidermal growth factor-like 1).
GN   CELSR3 OR CDHF11 OR FMI1 OR EGFL1 OR MEGF2.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20202599; PubMed=10716726;
RA   Wu Q., Maniatis T.;
RT   "Large exons encoding multiple ectodomains are a characteristic
RT   feature of protocadherin genes.";
RL   Proc. Natl. Acad. Sci. U.S.A. 97:3124-3129(2000).
RN   [2]
RP   SEQUENCE OF 1954-3312 FROM N.A.
```

```
RC   TISSUE=Brain;
RX   MEDLINE=98360089; PubMed=9693030;
RA   Nakayama M., Nakajima D., Nagase T., Nomura N., Seki N., Ohara O.;
RT   "Identification of high-molecular-weight proteins with multiple
RT   EGF-like motifs by motif-trap screening.";
RL   Genomics 51:27-34(1998).
CC   -!- FUNCTION: Receptor that may have an important role in cell/cell
CC       signaling during nervous system formation.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- SIMILARITY: Belongs to family 2 of G-protein coupled receptors.
CC   -!- SIMILARITY: Contains 9 cadherin domains.
CC   -!- SIMILARITY: Contains 8 EGF-like domains.
CC   -!- SIMILARITY: Contains 2 laminin G-like domains.
CC   -!- SIMILARITY: Contains 1 laminin EGF-like domain.
CC   -!- SIMILARITY: Contains 1 GPS domain.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF231023; AAF61929.1; -.
DR   EMBL; AB011536; BAA32464.1; -.
DR   HSSP; P00740; 1EDM.
DR   Genew; HGNC:3230; CELSR3.
DR   MIM; 604264; -.
DR   GO; GO:0005198; F:structural molecule activity; NAS.
DR   InterPro; IPR000152; Asx_hydroxyl_S.
DR   InterPro; IPR002126; Cadherin.
DR   InterPro; IPR008985; ConA_like_lec_gl.
DR   InterPro; IPR000742; EGF_2.
DR   InterPro; IPR001881; EGF_Ca.
DR   InterPro; IPR006209; EGF_like.
DR   InterPro; IPR000832; GPCR_secretin.
DR   InterPro; IPR001879; hormn_receptor.
DR   InterPro; IPR006210; IEGF.
DR   InterPro; IPR002049; Laminin_EGF.
DR   InterPro; IPR001791; Laminin_G.
DR   InterPro; IPR000203; PKD_cys_rich.
DR   Pfam; PF00002; 7tm_2; 1.
DR   Pfam; PF00028; cadherin; 9.
DR   Pfam; PF00008; EGF; 5.
DR   Pfam; PF01825; GPS; 1.
DR   Pfam; PF02793; HRM; 1.
DR   Pfam; PF00054; laminin_G; 2.
DR   PRINTS; PR00205; CADHERIN.
DR   PRINTS; PR00011; EGFLAMININ.
DR   PRINTS; PR00249; GPCRSECRETIN.
DR   SMART; SM00112; CA; 9.
DR   SMART; SM00181; EGF; 6.
DR   SMART; SM00303; GPS; 1.
DR   SMART; SM00008; HormR; 1.
DR   SMART; SM00282; LamG; 2.
DR   PROSITE; PS00010; ASX_HYDROXYL; 1.
DR   PROSITE; PS00232; CADHERIN_1; 7.
DR   PROSITE; PS50268; CADHERIN_2; 8.
DR   PROSITE; PS00022; EGF_1; 6.
DR   PROSITE; PS01186; EGF_2; 4.
DR   PROSITE; PS50026; EGF_3; 6.
DR   PROSITE; PS00649; G_PROTEIN_RECEP_F2_1; FALSE_NEG.
DR   PROSITE; PS00650; G_PROTEIN_RECEP_F2_2; FALSE_NEG.
DR   PROSITE; PS50227; G_PROTEIN_RECEP_F2_3; 1.
DR   PROSITE; PS50261; G_PROTEIN_RECEP_F2_4; 1.
DR   PROSITE; PS50221; GPS; 1.
DR   PROSITE; PS50025; LAM_G_DOMAIN; 2.
DR   PROSITE; PS01248; LAMININ_TYPE_EGF; 1.
KW   G-protein coupled receptor; Transmembrane; Glycoprotein;
KW   EGF-like domain; Calcium-binding; Laminin EGF-like domain; Repeat;
KW   Developmental protein; Hydroxylation; Signal.
FT   SIGNAL        1     32       POTENTIAL.
FT   CHAIN        33   3312       CADHERIN EGF LAG SEVEN-PASS G-TYPE
FT                                RECEPTOR 3.
FT   DOMAIN       33   2540       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   2541   2561       1 (POTENTIAL).
FT   DOMAIN     2562   2572       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM   2573   2593       2 (POTENTIAL).
FT   DOMAIN     2594   2601       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   2602   2622       3 (POTENTIAL).
FT   DOMAIN     2623   2643       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM   2644   2664       4 (POTENTIAL).
FT   DOMAIN     2665   2681       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   2682   2702       5 (POTENTIAL).
FT   DOMAIN     2703   2725       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM   2726   2746       6 (POTENTIAL).
FT   DOMAIN     2747   2753       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   2754   2774       7 (POTENTIAL).
FT   DOMAIN     2775   3312       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN      326    433       CADHERIN 1.
FT   DOMAIN      434    545       CADHERIN 2.
FT   DOMAIN      546    651       CADHERIN 3.
FT   DOMAIN      652    756       CADHERIN 4.
FT   DOMAIN      757    858       CADHERIN 5.
FT   DOMAIN      859    961       CADHERIN 6.
FT   DOMAIN      962   1067       CADHERIN 7.
FT   DOMAIN     1068   1169       CADHERIN 8.
FT   DOMAIN     1170   1265       CADHERIN 9.
FT   DOMAIN     1375   1433       EGF-LIKE 1, CALCIUM-BINDING.
FT   DOMAIN     1435   1471       EGF-LIKE 2, CALCIUM-BINDING.
FT   DOMAIN     1475   1514       EGF-LIKE 3, CALCIUM-BINDING.
FT   DOMAIN     1515   1719       LAMININ G-LIKE 1.
FT   DOMAIN     1722   1758       EGF-LIKE 4, CALCIUM-BINDING.
FT   DOMAIN     1764   1944       LAMININ G-LIKE 2.
FT   DOMAIN     1946   1982       EGF-LIKE 5, CALCIUM-BINDING.
FT   DOMAIN     1983   2020       EGF-LIKE 6, CALCIUM-BINDING.
FT   DOMAIN     2021   2053       EGF-LIKE 7, CALCIUM-BINDING.
FT   DOMAIN     2055   2090       EGF-LIKE 8, CALCIUM-BINDING.
FT   DOMAIN     2096   2131       LAMININ EGF-LIKE.
FT   DOMAIN     2477   2529       GPS.
FT   DISULFID   1379   1390       BY SIMILARITY.
FT   DISULFID   1384   1421       BY SIMILARITY.
FT   DISULFID   1423   1432       BY SIMILARITY.
FT   DISULFID   1439   1450       BY SIMILARITY.
FT   DISULFID   1444   1459       BY SIMILARITY.
FT   DISULFID   1461   1470       BY SIMILARITY.
FT   DISULFID   1479   1490       BY SIMILARITY.
FT   DISULFID   1484   1500       BY SIMILARITY.
FT   DISULFID   1502   1513       BY SIMILARITY.
FT   DISULFID   1726   1737       BY SIMILARITY.
FT   DISULFID   1731   1746       BY SIMILARITY.
FT   DISULFID   1748   1757       BY SIMILARITY.
FT   DISULFID   1950   1961       BY SIMILARITY.
FT   DISULFID   1955   1970       BY SIMILARITY.
FT   DISULFID   1972   1981       BY SIMILARITY.
FT   DISULFID   1985   1996       BY SIMILARITY.
FT   DISULFID   1990   2008       BY SIMILARITY.
FT   DISULFID   2010   2019       BY SIMILARITY.
FT   DISULFID   2027   2040       BY SIMILARITY.
FT   DISULFID   2042   2052       BY SIMILARITY.
FT   DISULFID   2059   2074       BY SIMILARITY.
FT   DISULFID   2061   2077       BY SIMILARITY.
FT   DISULFID   2079   2089       BY SIMILARITY.
FT   MOD_RES    1963   1963       HYDROXYLATION (POTENTIAL).
FT   CARBOHYD    632    632       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    847    847       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1182   1182       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1222   1222       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1317   1317       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1327   1327       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1649   1649       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1713   1713       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1770   1770       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   2053   2053       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   2177   2177       N-LINKED (GLCNAC. . .) (POTENTIAL).
```

```
FT   CARBOHYD   2196   2196       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   2386   2386       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   2474   2474       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   2506   2506       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CONFLICT   2158   2158       G -> GLRGAG (IN REF. 2).
SQ   SEQUENCE   3312 AA;  358251 MW;  BEC208703651A4A5 CRC64;

  Query Match             53.5%;  Score 46;  DB 1;  Length 3312;
  Best Local Similarity   61.5%;  Pred. No. 2e+02;
  Matches    8;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 DKGXXLPRPTPPR 13
              |:|  |||  |||
Db       3116 DRGSTLPRRQPPR 3128


RESULT 10
CLR3_RAT
ID   CLR3_RAT       STANDARD;      PRT;  3313 AA.
AC   O88278;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Cadherin EGF LAG seven-pass G-type receptor 3 precursor (Multiple
DE   epidermal growth factor-like domains 2).
GN   CELSR3 OR MEGF2.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Sprague-Dawley; TISSUE=Brain;
RX   MEDLINE=98360089; PubMed=9693030;
RA   Nakayama M., Nakajima D., Nagase T., Nomura N., Seki N., Ohara O.;
RT   "Identification of high-molecular-weight proteins with multiple
RT   EGF-like motifs by motif-trap screening.";
RL   Genomics 51:27-34(1998).
CC   -!- FUNCTION: Receptor that may have an important role in cell/cell
CC       signaling during nervous system formation.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- TISSUE SPECIFICITY: Expressed in the brain. Expressed in
CC       cerebrellum, olfactory bulb, cerebral cortex, hippocampus and
CC       brain stem.
CC   -!- SIMILARITY: Belongs to family 2 of G-protein coupled receptors.
CC   -!- SIMILARITY: Contains 9 cadherin domains.
CC   -!- SIMILARITY: Contains 8 EGF-like domains.
CC   -!- SIMILARITY: Contains 2 laminin G-like domains.
CC   -!- SIMILARITY: Contains 1 laminin EGF-like domain.
CC   -!- SIMILARITY: Contains 1 GPS domain.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AB011528; BAA32459.1; -.
DR   HSSP; P00740; 1EDM.
DR   InterPro; IPR000152; Asx_hydroxyl_S.
DR   InterPro; IPR002126; Cadherin.
DR   InterPro; IPR008985; ConA_like_lec_gl.
DR   InterPro; IPR000742; EGF_2.
DR   InterPro; IPR001881; EGF_Ca.
DR   InterPro; IPR006209; EGF_like.
DR   InterPro; IPR000832; GPCR_secretin.
DR   InterPro; IPR001879; hormn_receptor.
DR   InterPro; IPR006210; IEGF.
DR   InterPro; IPR002049; Laminin_EGF.
DR   InterPro; IPR001791; Laminin_G.
DR   InterPro; IPR000203; PKD_cys_rich.
DR   Pfam; PF00002; 7tm_2; 1.
DR   Pfam; PF00028; cadherin; 9.
DR   Pfam; PF00008; EGF; 5.
DR   Pfam; PF01825; GPS; 1.
DR   Pfam; PF02793; HRM; 1.
DR   Pfam; PF00054; laminin_G; 1.
DR   PRINTS; PR00205; CADHERIN.
DR   PRINTS; PR00011; EGFLAMININ.
DR   PRINTS; PR00249; GPCRSECRETIN.
DR   SMART; SM00112; CA; 9.
DR   SMART; SM00181; EGF; 6.
DR   SMART; SM00303; GPS; 1.
DR   SMART; SM00008; HormR; 1.
DR   SMART; SM00282; LamG; 2.
DR   PROSITE; PS00010; ASX_HYDROXYL; 1.
DR   PROSITE; PS00232; CADHERIN_1; 7.
DR   PROSITE; PS50268; CADHERIN_2; 8.
DR   PROSITE; PS00022; EGF_1; 6.
DR   PROSITE; PS01186; EGF_2; 4.
DR   PROSITE; PS50026; EGF_3; 6.
DR   PROSITE; PS00649; G_PROTEIN_RECEP_F2_1; FALSE_NEG.
DR   PROSITE; PS00650; G_PROTEIN_RECEP_F2_2; FALSE_NEG.
DR   PROSITE; PS50227; G_PROTEIN_RECEP_F2_3; 1.
DR   PROSITE; PS50261; G_PROTEIN_RECEP_F2_4; 1.
DR   PROSITE; PS50221; GPS; 1.
DR   PROSITE; PS50025; LAM_G_DOMAIN; 2.
DR   PROSITE; PS01248; LAMININ_TYPE_EGF; 1.
KW   G-protein coupled receptor; Transmembrane; Glycoprotein;
KW   EGF-like domain; Calcium-binding; Laminin EGF-like domain; Repeat;
KW   Developmental protein; Hydroxylation; Signal.
FT   SIGNAL        1     31       POTENTIAL.
FT   CHAIN        32   3313       CADHERIN EGF LAG SEVEN-PASS G-TYPE
FT                                RECEPTOR 3.
FT   DOMAIN       32   2538       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   2539   2559       1 (POTENTIAL).
FT   DOMAIN     2560   2570       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM   2571   2591       2 (POTENTIAL).
FT   DOMAIN     2592   2599       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   2600   2620       3 (POTENTIAL).
FT   DOMAIN     2621   2641       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM   2642   2662       4 (POTENTIAL).
FT   DOMAIN     2663   2679       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   2680   2700       5 (POTENTIAL).
FT   DOMAIN     2701   2724       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM   2725   2745       6 (POTENTIAL).
FT   DOMAIN     2746   2752       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   2753   2773       7 (POTENTIAL).
FT   DOMAIN     2774   3313       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN      317    424       CADHERIN 1.
FT   DOMAIN      425    536       CADHERIN 2.
FT   DOMAIN      537    642       CADHERIN 3.
FT   DOMAIN      643    747       CADHERIN 4.
FT   DOMAIN      748    849       CADHERIN 5.
FT   DOMAIN      850    952       CADHERIN 6.
FT   DOMAIN      953   1058       CADHERIN 7.
FT   DOMAIN     1059   1160       CADHERIN 8.
FT   DOMAIN     1161   1257       CADHERIN 9.
FT   DOMAIN     1366   1424       EGF-LIKE 1, CALCIUM-BINDING.
FT   DOMAIN     1426   1462       EGF-LIKE 2, CALCIUM-BINDING.
FT   DOMAIN     1466   1505       EGF-LIKE 3, CALCIUM-BINDING.
FT   DOMAIN     1506   1710       LAMININ G-LIKE 1.
FT   DOMAIN     1713   1749       EGF-LIKE 4, CALCIUM-BINDING.
FT   DOMAIN     1753   1935       LAMININ G-LIKE 2.
FT   DOMAIN     1937   1972       EGF-LIKE 5, CALCIUM-BINDING.
FT   DOMAIN     1973   2011       EGF-LIKE 6, CALCIUM-BINDING.
FT   DOMAIN     2012   2044       EGF-LIKE 7, CALCIUM-BINDING.
FT   DOMAIN     2046   2081       EGF-LIKE 8, CALCIUM-BINDING.
FT   DOMAIN     2087   2120       LAMININ EGF-LIKE.
FT   DOMAIN     2475   2527       GPS.
FT   DISULFID   1370   1381       BY SIMILARITY.
FT   DISULFID   1375   1412       BY SIMILARITY.
FT   DISULFID   1414   1423       BY SIMILARITY.
FT   DISULFID   1430   1441       BY SIMILARITY.
```

```
FT   DISULFID   1435   1450       BY SIMILARITY.
FT   DISULFID   1452   1461       BY SIMILARITY.
FT   DISULFID   1470   1481       BY SIMILARITY.
FT   DISULFID   1475   1491       BY SIMILARITY.
FT   DISULFID   1493   1504       BY SIMILARITY.
FT   DISULFID   1717   1728       BY SIMILARITY.
FT   DISULFID   1722   1737       BY SIMILARITY.
FT   DISULFID   1739   1748       BY SIMILARITY.
FT   DISULFID   1941   1952       BY SIMILARITY.
FT   DISULFID   1946   1961       BY SIMILARITY.
FT   DISULFID   1963   1972       BY SIMILARITY.
FT   DISULFID   1976   1987       BY SIMILARITY.
FT   DISULFID   1981   1999       BY SIMILARITY.
FT   DISULFID   2001   2010       BY SIMILARITY.
FT   DISULFID   2018   2031       BY SIMILARITY.
FT   DISULFID   2033   2043       BY SIMILARITY.
FT   DISULFID   2050   2065       BY SIMILARITY.
FT   DISULFID   2052   2068       BY SIMILARITY.
FT   DISULFID   2070   2080       BY SIMILARITY.
FT   MOD_RES    1954   1954       HYDROXYLATION (POTENTIAL).
FT   CARBOHYD    623    623       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    838    838       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1173   1173       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1213   1213       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1308   1308       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1318   1318       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1640   1640       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1704   1704       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1761   1761       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   2044   2044       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   2173   2173       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   2192   2192       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   2382   2382       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   2472   2472       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   2504   2504       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   3313 AA;  359348 MW;  B11DA09517288764 CRC64;

  Query Match             53.5%;  Score 46;  DB 1;  Length 3313;
  Best Local Similarity   61.5%;  Pred. No. 2e+02;
  Matches    8;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 DKGXXLPRPTPPR 13
              |:|  |||  |||
Db       3117 DRGSTLPRRQPPR 3129


RESULT 11
EXON_HSV2
ID   EXON_HSV2      STANDARD;      PRT;   620 AA.
AC   P06489; Q69352;
DT   01-JAN-1988 (Rel. 06, Created)
DT   01-JAN-1988 (Rel. 06, Last sequence update)
DT   15-JUL-1998 (Rel. 36, Last annotation update)
DE   Alkaline exonuclease (EC 3.1.11.-).
GN   UL12.
OS   Herpes simplex virus (type 2).
OC   Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC   Alphaherpesvirinae; Simplexvirus.
OX   NCBI_TaxID=10310;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=86144016; PubMed=3005609;
RA   Draper K.G., Devi-Rao G., Costa R.H., Blair E.D., Thompson R.L.,
RA   Wagner E.K.;
RT   "Characterization of the genes encoding herpes simplex virus type 1
RT   and type 2 alkaline exonucleases and overlapping proteins.";
RL   J. Virol. 57:1023-1036(1986).
CC   -!- SIMILARITY: Belongs to the herpesviruses alkaline exonuclease
CC       family.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M11854; AAA45834.1; -.
DR   EMBL; M11854; AAA45835.1; ALT_INIT.
DR   InterPro; IPR001616; Herpes_alk_exo.
DR   Pfam; PF01771; Herpes_alk_exo; 1.
DR   PRINTS; PR00924; ALKEXNUCLASE.
KW   Hydrolase; Nuclease; Exonuclease.
SQ   SEQUENCE   620 AA;  66199 MW;  3E4E89AC766414B7 CRC64;

  Query Match             52.3%;  Score 45;  DB 1;  Length 620;
  Best Local Similarity   87.5%;  Pred. No. 50;
  Matches    7;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          7 PRPTPPRP 14
              ||| ||||
Db         39 PRPAPPRP 46


RESULT 12
DIA1_MOUSE
ID   DIA1_MOUSE     STANDARD;      PRT;  1255 AA.
AC   O08808;
DT   15-JUL-1999 (Rel. 38, Created)
DT   15-JUL-1999 (Rel. 38, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   Diaphanous protein homolog 1 (Diaphanous-related formin 1) (DRF1)
DE   (mDIA1) (p140mDIA).
GN   DIAPH1 OR DIAP1.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=97357293; PubMed=9214622;
RA   Watanabe N., Madaule P., Reid T., Ishizaki T., Watanabe G.,
RA   Kakizuka A., Saito Y., Nakao K., Jockusch B.M., Narumiya S.;
RT   "p140mDia, a mammalian homolog of Drosophila diaphanous, is a target
RT   protein for Rho small GTPase and is a ligand for profilin.";
RL   EMBO J. 16:3044-3056(1997).
RN   [2]
RP   FUNCTION.
RX   MEDLINE=20142655; PubMed=10678165;
RA   Tominaga T., Sahai E., Chardin P., McCormick F., Courtneidge S.A.,
RA   Alberts A.S.;
RT   "Diaphanous-related formins bridge Rho GTPase and Src tyrosine kinase
RT   signaling.";
RL   Mol. Cell 5:13-25(2000).
CC   -!- FUNCTION: Binds to GTP-bound form of Rho and to profilin. Acts in
CC       a Rho-dependent manner to recruit profilin to the membrane, where
CC       it promotes actin polymerization. It is required for cytokinesis,
CC       stress fiber formation, and transcriptional activation of the
CC       serum response factor. DFR proteins couple Rho and Src tyrosine
CC       kinase during signaling and the regulation of actin dynamics.
CC   -!- SUBCELLULAR LOCATION: MEMBRANE RUFFLES, ESPECIALLY AT THE TIP OF
CC       RUFFLES, OF MOTILE CELLS.
CC   -!- TISSUE SPECIFICITY: Ubiquitous.
CC   -!- DOMAIN: DRFs are regulated by intramolecular GBD-DAD binding where
CC       Rho-GTP activates the DRFs by disrupting the GBD-DAD interaction.
CC   -!- SIMILARITY: Contains 1 GTPase-binding (GBD) domain.
CC   -!- SIMILARITY: Contains 1 Formin homology 1 (FH1) domain.
CC   -!- SIMILARITY: Contains 1 Formin homology 2 (FH2) domain.
CC   -!- SIMILARITY: Contains 1 Formin homology 3 (FH3) domain.
CC   -!- SIMILARITY: Contains 1 DRF autoregulatory (DAD) domain.
CC   -!- SIMILARITY: Belongs to the formin homology family. Diaphanous
CC       subfamily.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
```

```
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; U96963; AAC53280.1; -.
DR   PIR; T31065; T31065.
DR   MGD; MGI:1194490; Diap1.
DR   GO; GO:0005515; F:protein binding; IPI.
DR   InterPro; IPR003104; FH2.
DR   Pfam; PF02181; FH2; 1.
DR   SMART; SM00498; FH2; 1.
KW   Coiled coil; Repeat.
FT   DOMAIN      460    562       COILED COIL (POTENTIAL).
FT   DOMAIN       63    260       GBD.
FT   DOMAIN      157    457       FH3.
FT   DOMAIN      586    747       FH1 (PRO-RICH).
FT   DOMAIN      752   1197       FH2.
FT   DOMAIN     1027   1179       COILED COIL (POTENTIAL).
FT   DOMAIN     1180   1194       DAD.
FT   DOMAIN     1196   1199       ARG/LYS-RICH (BASIC).
SQ   SEQUENCE   1255 AA;  139343 MW;  09404164873CA7C1 CRC64;

  Query Match             52.3%;  Score 45;  DB 1;  Length 1255;
  Best Local Similarity   46.7%;  Pred. No. 1e+02;
  Matches    7;  Conservative    2;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 DKGXXLPRPTPPRPI 15
              |  |   :| | || |:
Db        580 DSGTVIPPPPPPPPL 594


RESULT 13
STNB_DROME
ID   STNB_DROME     STANDARD;      PRT;  1262 AA.
AC   Q24212; Q9W5J3;
DT   10-OCT-2003 (Rel. 42, Created)
DT   10-OCT-2003 (Rel. 42, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   Stoned B protein (StonedB) (Snt-B).
GN   STNB OR CG12473/CG40302.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Oregon-R; TISSUE=CNS;
RX   MEDLINE=97001127; PubMed=8844157;
RA   Andrews J., Smith M., Merakovsky J., Coulson M., Hannan F.,
RA   Kelly L.E.;
RT   "The stoned locus of Drosophila melanogaster produces a dicistronic
RT   transcript and encodes two distinct polypeptides.";
RL   Genetics 143:1699-1711(1996).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Berkeley;
RX   MEDLINE=20196006; PubMed=10731132;
RA   Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA   Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA   George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA   Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA   Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA   Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA   Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA   Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA   Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA   Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA   Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA   Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA   de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA   Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA   Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA   Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA   Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA   Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
RA   Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA   Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA   Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA   Lasko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
RA   Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA   Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA   Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA   Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA   Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA   Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA   Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA   Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA   Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA   Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA   Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA   Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA   Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA   Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT   "The genome sequence of Drosophila melanogaster.";
RL   Science 287:2185-2195(2000).
RN   [3]
RP   FUNCTION, SUBCELLULAR LOCATION, AND DEVELOPMENTAL STAGE.
RX   MEDLINE=99337940; PubMed=10407025;
RA   Fergestad T., Davis W.S., Broadie K.;
RT   "The stoned proteins regulate synaptic vesicle recycling in the
RT   presynaptic terminal.";
RL   J. Neurosci. 19:5847-5860(1999).
RN   [4]
RP   INTERACTION WITH SYT.
RX   MEDLINE=20524362; PubMed=11069931;
RA   Phillips A.M., Smith M., Ramaswami M., Kelly L.E.;
RT   "The products of the Drosophila stoned locus interact with synaptic
RT   vesicles via synaptotagmin.";
RL   J. Neurosci. 20:8254-8261(2000).
RN   [5]
RP   FUNCTION, AND SUBCELLULAR LOCATION.
RX   MEDLINE=21114085; PubMed=11160392;
RA   Fergestad T., Broadie K.;
RT   "Interaction of stoned and synaptotagmin in synaptic vesicle
RT   endocytosis.";
RL   J. Neurosci. 21:1218-1227(2001).
RN   [6]
RP   FUNCTION.
RX   MEDLINE=21212245; PubMed=11312288;
RA   Stimson D.T., Estes P.S., Rao S., Krishnan K.S., Kelly L.E.,
RA   Ramaswami M.;
RT   "Drosophila stoned proteins regulate the rate and fidelity of synaptic
RT   vesicle internalization.";
RL   J. Neurosci. 21:3034-3044(2001).
RN   [7]
RP   RNA EDITING OF POSITION 1186.
RX   MEDLINE=22789647; PubMed=12907802;
RA   Hoopengardner B., Bhalla T., Staber C., Reenan R.;
RT   "Nervous system targets of RNA editing identified by comparative
RT   genomics.";
RL   Science 301:832-836(2003).
CC   -!- FUNCTION: Adapter protein involved in endocytic recycling of
CC       synaptic vesicles membranes. May act by mediating the retrieval of
CC       synaptotagmin protein Syt from the plasma membrane, thereby
CC       facilitating the internalization of multiple synaptic vesicles
CC       from the plasma membrane.
CC   -!- SUBUNIT: Interacts with the second C2 domain of Syt.
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic; colocalizes with synaptic
CC       vesicle pools. Colocalizes with the endocytic network within
CC       synaptic boutons.
CC   -!- DEVELOPMENTAL STAGE: Present at synaptic connections both in the
CC       CNS and in neuromuscular junctions in the mature embryo (20-22h)
CC       and throughout larval development. In the third instar larva, it
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: March 18, 2004, 05:58:16 ; Search time 39 Seconds
(without alignments)
119.518 Million cell updates/sec

Title: US-09-980-804-1
Perfect score: 86
Sequence: 1 DKGXXLRRPTPPRPIYXX 18

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1049977 seqs, 258955339 residues

Total number of hits satisfying chosen parameters: 1049977

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	81	94.2	20	14	US-10-181-654-3	Sequence 3, Appli
2	81	94.2	21	14	US-10-181-654-12	Sequence 12, Appl
3	81	94.2	21	14	US-10-181-654-25	Sequence 25, Appl
4	80	93.0	18	14	US-10-181-654-9	Sequence 9, Appli
5	74	86.0	18	14	US-10-181-654-36	Sequence 36, Appl
6	52	60.5	574	12	US-10-424-599-283004	Sequence 283004,
7	51	59.3	333	12	US-10-425-114-61578	Sequence 61578, A
8	51	59.3	383	12	US-10-425-114-53657	Sequence 53657, A
9	51	59.3	398	12	US-10-425-114-46621	Sequence 46621, A
10	51	59.3	401	12	US-10-425-114-40384	Sequence 40384, A
11	50	58.1	79	12	US-10-424-599-197137	Sequence 197137,
12	50	58.1	286	12	US-10-425-114-55260	Sequence 55260, A
13	50	58.1	290	12	US-10-425-114-67727	Sequence 67727, A
14	50	58.1	309	12	US-10-425-114-41158	Sequence 41158, A
15	48	55.8	197	12	US-10-425-114-42546	Sequence 42546, A
16	47	54.7	53	12	US-10-424-599-274473	Sequence 274473,
17	47	54.7	176	9	US-09-953-342-25	Sequence 25, Appl
18	47	54.7	304	12	US-10-425-114-42016	Sequence 42016, A
19	47	54.7	392	14	US-10-156-761-11324	Sequence 11324, A
20	47	54.7	1071	12	US-10-188-248-24	Sequence 24, Appl
21	47	54.7	1126	15	US-10-108-260A-3665	Sequence 3665, Ap
22	46	53.5	20	14	US-10-181-654-7	Sequence 7, Appli
23	46	53.5	89	12	US-10-424-599-248850	Sequence 248850,
24	46	53.5	107	12	US-10-424-599-169442	Sequence 169442,
25	46	53.5	487	14	US-10-224-999A-3465	Sequence 3465, Ap
26	46	53.5	2803	12	US-10-415-187-5	Sequence 5, Appli
27	46	53.5	3298	14	US-10-149-819-21	Sequence 21, Appl
28	46	53.5	3301	16	US-10-038-854-68	Sequence 68, Appl
29	46	53.5	3312	14	US-10-225-567A-656	Sequence 656, App
30	46	53.5	3312	16	US-10-038-854-67	Sequence 67, Appl
31	46	53.5	3313	9	US-09-737-149-29	Sequence 29, Appl
32	46	53.5	3313	16	US-10-038-854-69	Sequence 69, Appl
33	46	53.5	4115	16	US-10-038-854-4	Sequence 4, Appli
34	45	52.3	57	12	US-10-424-599-212361	Sequence 212361,
35	45	52.3	153	12	US-10-425-114-53570	Sequence 53570, A
36	45	52.3	190	12	US-10-424-599-166807	Sequence 166807,
37	45	52.3	199	14	US-10-034-934-125	Sequence 125, App
38	45	52.3	228	12	US-10-424-599-246968	Sequence 246968,
39	45	52.3	244	12	US-10-424-599-210656	Sequence 210656,
40	45	52.3	309	12	US-10-425-114-60031	Sequence 60031, A
41	45	52.3	358	12	US-10-425-114-45552	Sequence 45552, A
42	45	52.3	358	12	US-10-425-114-57738	Sequence 57738, A
43	45	52.3	434	14	US-10-180-375-124	Sequence 124, App
44	45	52.3	489	12	US-10-425-114-50041	Sequence 50041, A
45	44	51.2	11	14	US-10-161-791-294	Sequence 294, App

ALIGNMENTS

RESULT 1
US-10-181-654-3
; Sequence 3, Application US/10181654
; Publication No. US20030108957A1
; GENERAL INFORMATION:
; APPLICANT: The Wistar Institute of Anatomy and Biology
; APPLICANT: Creighton University
; APPLICANT: Otvos, Laszlo
; APPLICANT: Blaszczyk-Thurin, Magdalena
; APPLICANT: Rogers, Mark
; APPLICANT: Lovas, Sandor
; TITLE OF INVENTION: Biocidal Molecules, Macromolecular Targets and Methods of Product
; TITLE OF INVENTION: Use
; FILE REFERENCE: WST94BPCT
; CURRENT APPLICATION NUMBER: US/10/181,654
; CURRENT FILING DATE: 2002-07-19
; PRIOR APPLICATION NUMBER: US 60/177,565
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: US 60/237,599
; PRIOR FILING DATE: 2000-10-03
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 3
; LENGTH: 20
; TYPE: PRT
; ORGANISM: P. apterus
US-10-181-654-3

Query Match 94.2%; Score 81; DB 14; Length 20;
Best Local Similarity 87.5%; Pred. No. 0.0017;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 DKGXXLPRPTPPRPIY 16
||| |||||||||||
Db 2 DKGSYLPRPTPPRPIY 17

RESULT 2

```
US-10-181-654-12
; Sequence 12, Application US/10181654
; Publication No. US20030108957A1
; GENERAL INFORMATION:
;  APPLICANT: The Wistar Institute of Anatomy and Biology
;  APPLICANT:  Creighton University
;  APPLICANT:  Otvos, Laszlo
;  APPLICANT:  Blaszczyk-Thurin, Magdalena
;  APPLICANT:  Rogers, Mark
;  APPLICANT:  Lovas, Sandor
;  TITLE OF INVENTION: Biocidal Molecules, Macromolecular Targets and Methods of Product
;  TITLE OF INVENTION:  Use
;  FILE REFERENCE: WST94BPCT
;  CURRENT APPLICATION NUMBER: US/10/181,654
;  CURRENT FILING DATE:  2002-07-19
;  PRIOR APPLICATION NUMBER: US 60/177,565
;  PRIOR FILING DATE: 2000-01-21
;  PRIOR APPLICATION NUMBER: US 60/237,599
;  PRIOR FILING DATE: 2000-10-03
;  NUMBER OF SEQ ID NOS: 36
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 12
;   LENGTH: 21
;   TYPE: PRT
;   ORGANISM: biotin-K-pyrrhocoricin
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(1)
;   OTHER INFORMATION: biotin is attached to Lys in position 1
US-10-181-654-12

  Query Match             94.2%;  Score 81;  DB 14;  Length 21;
  Best Local Similarity   87.5%;  Pred. No. 0.0018;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 DKGXXLPRPTPPRPIY 16
              |||  ||||||||||||
Db          3 DKGSYLPRPTPPRPIY 18


RESULT 3
US-10-181-654-25
; Sequence 25, Application US/10181654
; Publication No. US20030108957A1
; GENERAL INFORMATION:
;  APPLICANT: The Wistar Institute of Anatomy and Biology
;  APPLICANT:  Creighton University
;  APPLICANT:  Otvos, Laszlo
;  APPLICANT:  Blaszczyk-Thurin, Magdalena
;  APPLICANT:  Rogers, Mark
;  APPLICANT:  Lovas, Sandor
;  TITLE OF INVENTION: Biocidal Molecules, Macromolecular Targets and Methods of Product
;  TITLE OF INVENTION:  Use
;  FILE REFERENCE: WST94BPCT
;  CURRENT APPLICATION NUMBER: US/10/181,654
;  CURRENT FILING DATE:  2002-07-19
;  PRIOR APPLICATION NUMBER: US 60/177,565
;  PRIOR FILING DATE: 2000-01-21
;  PRIOR APPLICATION NUMBER: US 60/237,599
;  PRIOR FILING DATE: 2000-10-03
;  NUMBER OF SEQ ID NOS: 36
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 25
;   LENGTH: 21
;   TYPE: PRT
;   ORGANISM: fluorescein-K pyrrhocoricin
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(1)
;   OTHER INFORMATION: fluorescein is attached to Lys in position 1
US-10-181-654-25

  Query Match             94.2%;  Score 81;  DB 14;  Length 21;
  Best Local Similarity   87.5%;  Pred. No. 0.0018;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 DKGXXLPRPTPPRPIY 16
              |||  ||||||||||||
Db          3 DKGSYLPRPTPPRPIY 18


RESULT 4
US-10-181-654-9
; Sequence 9, Application US/10181654
; Publication No. US20030108957A1
; GENERAL INFORMATION:
;  APPLICANT: The Wistar Institute of Anatomy and Biology
;  APPLICANT:  Creighton University
;  APPLICANT:  Otvos, Laszlo
;  APPLICANT:  Blaszczyk-Thurin, Magdalena
;  APPLICANT:  Rogers, Mark
;  APPLICANT:  Lovas, Sandor
;  TITLE OF INVENTION: Biocidal Molecules, Macromolecular Targets and Methods of Product
;  TITLE OF INVENTION:  Use
;  FILE REFERENCE: WST94BPCT
;  CURRENT APPLICATION NUMBER: US/10/181,654
;  CURRENT FILING DATE:  2002-07-19
;  PRIOR APPLICATION NUMBER: US 60/177,565
;  PRIOR FILING DATE: 2000-01-21
;  PRIOR APPLICATION NUMBER: US 60/237,599
;  PRIOR FILING DATE: 2000-10-03
;  NUMBER OF SEQ ID NOS: 36
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 9
;   LENGTH: 18
;   TYPE: PRT
;   ORGANISM: modified pyrrhocoricin peptide
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(1)
;   OTHER INFORMATION: A moiety having a net positive charge is attached to Asp
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (4)..(4)
;   OTHER INFORMATION: can be any amino acid
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (5)..(5)
;   OTHER INFORMATION: can be any amino acid
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (17)..(17)
;   OTHER INFORMATION: can be any amino acid
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (18)..(18)
;   OTHER INFORMATION: can be any amino acid and can have attached any of the groups con
;   OTHER INFORMATION: s
US-10-181-654-9

  Query Match             93.0%;  Score 80;  DB 14;  Length 18;
  Best Local Similarity   100.0%;  Pred. No. 0.0021;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 DKGXXLPRPTPPRPIY 16
              ||||||||||||||||
Db          1 DKGXXLPRPTPPRPIY 16


RESULT 5
US-10-181-654-36
; Sequence 36, Application US/10181654
; Publication No. US20030108957A1
; GENERAL INFORMATION:
;  APPLICANT: The Wistar Institute of Anatomy and Biology
;  APPLICANT:  Creighton University
```

```
;  APPLICANT:  Otvos, Laszlo
;  APPLICANT:  Blaszczyk-Thurin, Magdalena
;  APPLICANT:  Rogers, Mark
;  APPLICANT:  Lovas, Sandor
;  TITLE OF INVENTION: Biocidal Molecules, Macromolecular Targets and Methods of Product
;  TITLE OF INVENTION:  Use
;  FILE REFERENCE: WST94BPCT
;  CURRENT APPLICATION NUMBER: US/10/181,654
;  CURRENT FILING DATE:  2002-07-19
;  PRIOR APPLICATION NUMBER: US 60/177,565
;  PRIOR FILING DATE: 2000-01-21
;  PRIOR APPLICATION NUMBER: US 60/237,599
;  PRIOR FILING DATE: 2000-10-03
;  NUMBER OF SEQ ID NOS: 36
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 36
;   LENGTH: 18
;   TYPE: PRT
;   ORGANISM: modification of Pyrrhocoricin
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(1)
;   OTHER INFORMATION: Asp in position 1 is modified by a 1-aminocyclo-hexane carboxylic
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (18)..(18)
;   OTHER INFORMATION: Arg in position 18 is modified by an amino linker moiety
US-10-181-654-36

  Query Match           86.0%;  Score 74;  DB 14;  Length 18;
  Best Local Similarity   81.2%;  Pred. No. 0.011;
  Matches   13;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 DKGXXLPRPTPPRPIY 16
              | |  ||||||||||||
Db          1 DLGSYLPRPTPPRPIY 16


RESULT 6
US-10-424-599-283004
; Sequence 283004, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
;  APPLICANT: La Rosa Thomas J
;  APPLICANT:  Kovalic David K
;  APPLICANT:  Zhou Yihua
;  APPLICANT:  Cao Yongwei
;  TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants and Uses Thereof for Plant Improvement
;  FILE REFERENCE: 38-21(53223)B
;  CURRENT APPLICATION NUMBER: US/10/424,599
;  CURRENT FILING DATE:  2003-04-28
;  NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 283004
;   LENGTH: 574
;   TYPE: PRT
;   ORGANISM: Glycine max
;   FEATURE:
;   NAME/KEY: unsure
;   LOCATION: (1)..(574)
;   OTHER INFORMATION: unsure at all Xaa locations
;   FEATURE:
;   OTHER INFORMATION: Clone ID: PAT_MRT3847_97575C.1.pep
US-10-424-599-283004

  Query Match           60.5%;  Score 52;  DB 12;  Length 574;
  Best Local Similarity   80.0%;  Pred. No. 1.1e+02;
  Matches    8;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          7 PRPTPPRPIY 16
              |||:|||| |
Db        545 PRPSPPRPAY 554


RESULT 7
US-10-425-114-61578
; Sequence 61578, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
;  APPLICANT: Liu, Jingdong
;  APPLICANT:  Zhou, Yihua
;  APPLICANT:  Kovalic, David K.
;  APPLICANT:  Screen, Steven E
;  APPLICANT:  Tabaska, Jack E
;  APPLICANT:  Cao, Yongwei
;  TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants and Uses Thereof for Plant Improvement
;  FILE REFERENCE: 38-21(53313)B
;  CURRENT APPLICATION NUMBER: US/10/425,114
;  CURRENT FILING DATE:  2003-04-28
;  NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 61578
;   LENGTH: 333
;   TYPE: PRT
;   ORGANISM: Zea mays
;   FEATURE:
;   OTHER INFORMATION: Clone ID: LIB3069-035-G11_FLI.pep
US-10-425-114-61578

  Query Match           59.3%;  Score 51;  DB 12;  Length 333;
  Best Local Similarity   88.9%;  Pred. No. 89;
  Matches    8;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          6 LPRPTPPRP 14
              :||||||||
Db         64 VPRPTPPRP 72


RESULT 8
US-10-425-114-53657
; Sequence 53657, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
;  APPLICANT: Liu, Jingdong
;  APPLICANT:  Zhou, Yihua
;  APPLICANT:  Kovalic, David K.
;  APPLICANT:  Screen, Steven E
;  APPLICANT:  Tabaska, Jack E
;  APPLICANT:  Cao, Yongwei
;  TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants and Uses Thereof for Plant Improvement
;  FILE REFERENCE: 38-21(53313)B
;  CURRENT APPLICATION NUMBER: US/10/425,114
;  CURRENT FILING DATE:  2003-04-28
;  NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 53657
;   LENGTH: 383
;   TYPE: PRT
;   ORGANISM: Zea mays
;   FEATURE:
;   OTHER INFORMATION: Clone ID: LIB3150-064-E8_FLI.pep
US-10-425-114-53657

  Query Match           59.3%;  Score 51;  DB 12;  Length 383;
  Best Local Similarity   88.9%;  Pred. No. 1e+02;
  Matches    8;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          6 LPRPTPPRP 14
              :||||||||
Db        114 VPRPTPPRP 122


RESULT 9
US-10-425-114-46621
; Sequence 46621, Application US/10425114
; Publication No. US20040034888A1
```

; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 46621
; LENGTH: 398
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: 700456526_FLI.pep
US-10-425-114-46621

Query Match 59.3%; Score 51; DB 12; Length 398;
Best Local Similarity 88.9%; Pred. No. 1e+02;
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 6 LPRPTPPRP 14
:|||||||||
Db 186 VPRPTPPRP 194

RESULT 10
US-10-425-114-40384
; Sequence 40384, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 40384
; LENGTH: 401
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB143-013-C7_FLI.pep
US-10-425-114-40384

Query Match 59.3%; Score 51; DB 12; Length 401;
Best Local Similarity 88.9%; Pred. No. 1.1e+02;
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 6 LPRPTPPRP 14
:|||||||||
Db 189 VPRPTPPRP 197

RESULT 11
US-10-424-599-197137
; Sequence 197137, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 197137
; LENGTH: 79
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(79)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_2003C.1.pep
US-10-424-599-197137

Query Match 58.1%; Score 50; DB 12; Length 79;
Best Local Similarity 100.0%; Pred. No. 33;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 7 PRPTPPRP 14
||||||||
Db 36 PRPTPPRP 43

RESULT 12
US-10-425-114-55260
; Sequence 55260, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 55260
; LENGTH: 286
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: 701169290_FLI.pep
US-10-425-114-55260

Query Match 58.1%; Score 50; DB 12; Length 286;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 7 PRPTPPRP 14
||||||||
Db 173 PRPTPPRP 180

RESULT 13
US-10-425-114-67727
; Sequence 67727, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With

```
;  TITLE OF INVENTION:  Plants and Uses Thereof for Plant Improvement
;  FILE REFERENCE: 38-21(53313)B
;  CURRENT APPLICATION NUMBER: US/10/425,114
;  CURRENT FILING DATE:  2003-04-28
;  NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 67727
;   LENGTH: 290
;   TYPE: PRT
;   ORGANISM: Zea mays
;   FEATURE:
;   OTHER INFORMATION: Clone ID: LIB3597-046-D9_FLI.pep
US-10-425-114-67727

  Query Match           58.1%;  Score 50;  DB 12;  Length 290;
  Best Local Similarity  100.0%;  Pred. No. 1e+02;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          7 PRPTPPRP 14
              ||||||||
Db        173 PRPTPPRP 180


RESULT 14
US-10-425-114-41158
; Sequence 41158, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
;  APPLICANT: Liu, Jingdong
;  APPLICANT:  Zhou, Yihua
;  APPLICANT:  Kovalic, David K.
;  APPLICANT:  Screen, Steven E
;  APPLICANT:  Tabaska, Jack E
;  APPLICANT:  Cao, Yongwei
;  TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants and Uses Thereof for Plant Improvement
;  FILE REFERENCE: 38-21(53313)B
;  CURRENT APPLICATION NUMBER: US/10/425,114
;  CURRENT FILING DATE:  2003-04-28
;  NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 41158
;   LENGTH: 309
;   TYPE: PRT
;   ORGANISM: Zea mays
;   FEATURE:
;   OTHER INFORMATION: Clone ID: LIB3067-007-H4_FLI.pep
US-10-425-114-41158

  Query Match           58.1%;  Score 50;  DB 12;  Length 309;
  Best Local Similarity  100.0%;  Pred. No. 1.1e+02;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          7 PRPTPPRP 14
              ||||||||
Db        173 PRPTPPRP 180


RESULT 15
US-10-425-114-42546
; Sequence 42546, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
;  APPLICANT: Liu, Jingdong
;  APPLICANT:  Zhou, Yihua
;  APPLICANT:  Kovalic, David K.
;  APPLICANT:  Screen, Steven E
;  APPLICANT:  Tabaska, Jack E
;  APPLICANT:  Cao, Yongwei
;  TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants and Uses Thereof for Plant Improvement
;  FILE REFERENCE: 38-21(53313)B
;  CURRENT APPLICATION NUMBER: US/10/425,114
;  CURRENT FILING DATE:  2003-04-28
;  NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 42546
;   LENGTH: 197
;   TYPE: PRT
;   ORGANISM: Zea mays
;   FEATURE:
;   OTHER INFORMATION: Clone ID: 700223942_FLI.pep
US-10-425-114-42546

  Query Match           55.8%;  Score 48;  DB 12;  Length 197;
  Best Local Similarity  88.9%;  Pred. No. 1.3e+02;
  Matches    8;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          6 LPRPTPPRP 14
              |||| ||||
Db         54 LPRPPPPRP 62


Search completed: March 18, 2004, 06:07:21
Job time : 40 secs
```